

# Dynamic Conjunctive Queries\*

Thomas Zeume<sup>†</sup> and Thomas Schwentick<sup>‡</sup>

TU Dortmund University

April 6, 2017

## Abstract

The article investigates classes of queries maintainable by conjunctive queries (CQs) and their extensions and restrictions in the dynamic complexity framework of Patnaik and Immerman. Starting from the basic language of quantifier-free conjunctions of positive atoms, it studies the impact of additional operators and features — such as union, atomic negation and quantification — on the dynamic expressiveness, for the standard semantics as well as for  $\Delta$ -semantics.

Although many different combinations of these features are possible, they basically yield five important fragments for the standard semantics, characterized by the addition of (1) arbitrary quantification and atomic negation, (2) existential quantification and atomic negation, (3) existential quantification, (4) atomic negation (but no quantification)), and by (5) no addition to the basic language at all. While fragments (3), (4) and (5) can be separated, it remains open whether fragments (1), (2) and (3) are actually different. The fragments arising from  $\Delta$ -semantics are also subsumed by the standard fragments (1), (2) and (4). The main fragments of DYNFO that had been studied in previous work, DYNQF and DYNPROP, characterized by quantifier-free update programs with or without auxiliary functions, respectively, also fit into this hierarchy: DYNPROP coincides with fragment (4) and DYNQF is strictly above fragment (4) and within fragment (3).

As a further result, all (statically) FO-definable queries are captured by fragment (2) and a complete characterization of these queries in terms of non-recursive dynamic  $\exists^1$ FO-programs is given.

---

\*An extended abstract of this work appeared in the proceedings of the conference International Conference on Database Theory 2014 (ICDT 2014) [20]. Some preliminary results appeared in the proceedings of the conference Mathematical Foundations of Computer Science 2013 (MFCS 2013)[19]. Both authors acknowledge the financial support by DFG grant SCHW 678/6-1.

<sup>†</sup>thomas.zeume@cs.tu-dortmund.de

<sup>‡</sup>thomas.schwentick@tu-dortmund.de

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Dynamic setting</b>	<b>7</b>
<b>3</b>	<b>Dynamic conjunctive queries</b>	<b>9</b>
3.1	Collapse results . . . . .	10
3.2	Separation results . . . . .	21
<b>4</b>	<b><math>\Delta</math>-semantics</b>	<b>26</b>
<b>5</b>	<b>A dynamic characterization of first-order logic</b>	<b>34</b>
<b>6</b>	<b>Discussion and Future Work</b>	<b>38</b>

## 1 Introduction

The re-evaluation of a fixed query after a modification to a huge database can be a time-consuming process; in particular when it is performed from scratch. For this reason previously computed information such as the old query result and (possibly) other auxiliary information is often reused in order to speed up the process. This maintenance of query results has attracted lots of attention over the last decades of database related research, in particular in the related field of view maintenance. For relational databases algorithmic (see e.g. [17, 11]) and declarative approaches (see e.g. [7, 4, 16]) have been studied.

Here, we continue the study of the declarative approach where query results are maintained by queries from some query language. More precisely, for a relational database subject to change, auxiliary relations are maintained with the intention to help answering a query  $Q$ . When a modification to the database, i.e. an insertion or deletion of a tuple, occurs, every auxiliary relation is updated through a first-order query that can refer to both, the database and the auxiliary relations.

One possible formalization of this approach is the descriptive dynamic complexity framework (short: dynamic complexity) by Patnaik and Immerman [16]. In their framework, the class DYNFO contains of all queries maintainable through first-order updates (and thus also in the core of SQL). This is the formalization used in this work.

Shortly before the work of Patnaik and Immerman, the declarative approach was independently formalized in a very similar way by Dong, Su and Topor [7, 4]. For a discussion of the differences of both formalizations we refer to the later discussion of the choice of the precise setting.

The class DYNFO is quite powerful. Many queries inexpressible in (static) first-order logic, such as the transitive closure query on undirected graphs [16]

and the word problem for context-free languages [9], can be maintained in DYNFO. There are no general inexpressibility results for DYNFO at all<sup>1</sup>.

Towards a deeper understanding of the dynamic maintainability of queries, two main restrictions of DYNFO have been explored in the literature. Dong and Su started the study of restricted auxiliary relations [6] and obtained inexpressibility results for unary auxiliary relations. On the other hand, Hesse started the exploration of syntactic fragments of DYNFO, such as the one obtained by disallowing quantification in update formulas [13]. Inexpressibility results for this particular fragment have been obtained in [9].

In this work, we investigate classes of queries maintainable by conjunctive queries and extensions thereof, thus we are following the approach of Hesse.

Conjunctive queries (CQs), that is, in terms of logic, existential first-order queries whose quantifier-free part is a conjunction of atoms, are one of the most investigated query languages. Starting with Chandra and Merlin [1], who analyzed conjunctive queries for relational databases, those queries have been studied for almost every emerging new database model. Usually also the extension by unions (UCQs), by negations (CQ<sup>-</sup>s) as well as by both unions and negations (UCQ<sup>-</sup>s or, equivalently,  $\exists^*\text{FO}$ ) have been studied. It is folklore that all those classes are distinct for relational databases.

In this work we aim at the following goals.

**Goal 1.** *Understand the relative expressiveness of different extensions and restrictions of dynamic conjunctive queries in the dynamic setting, as well as their ability to maintain queries from (stronger) static query classes.*

For extensions, we add negation and/or disjunction to conjunctive queries, and for restrictions we disallow quantification. We further also consider universal quantification in place of existential quantification.

As for the relationship to static classes, it is interesting to understand whether larger static classes  $\mathcal{C}$  can be captured by dynamic classes  $\text{DYN}\mathcal{C}'$ , for weaker  $\mathcal{C}'$ . Up to now only two such results were known, namely, that MSO can be characterized by quantifier-free DYNFO on strings and that, on general structures,  $\exists^*\text{FO}$  is captured by DYNQF, the quantifier-free fragment of DYNFO with auxiliary relations *and* auxiliary functions [9].

In both dynamic complexity frameworks auxiliary relations are always explicitly defined as a whole after a modification. However, in the context of query re-evaluation, it is often convenient to express the new state of an auxiliary relation  $R$  in terms of the current relation and some “Delta”, that is, by specifying a tuple set  $R^+$  to be added to  $R$  and a tuple set  $R^-$  to be removed from  $R$ . We refer to the former semantics as *absolute semantics* and to the latter as  $\Delta$ -*semantics*. Obviously, the choice of the semantics does not affect the expressiveness of an update language that is closed under Boolean operations. However, most of the update languages in this paper lack some Boolean closure properties.

---

<sup>1</sup>Except for the trivial ones due to the fact that queries maintainable in DYNFO can be computed in polynomial time.

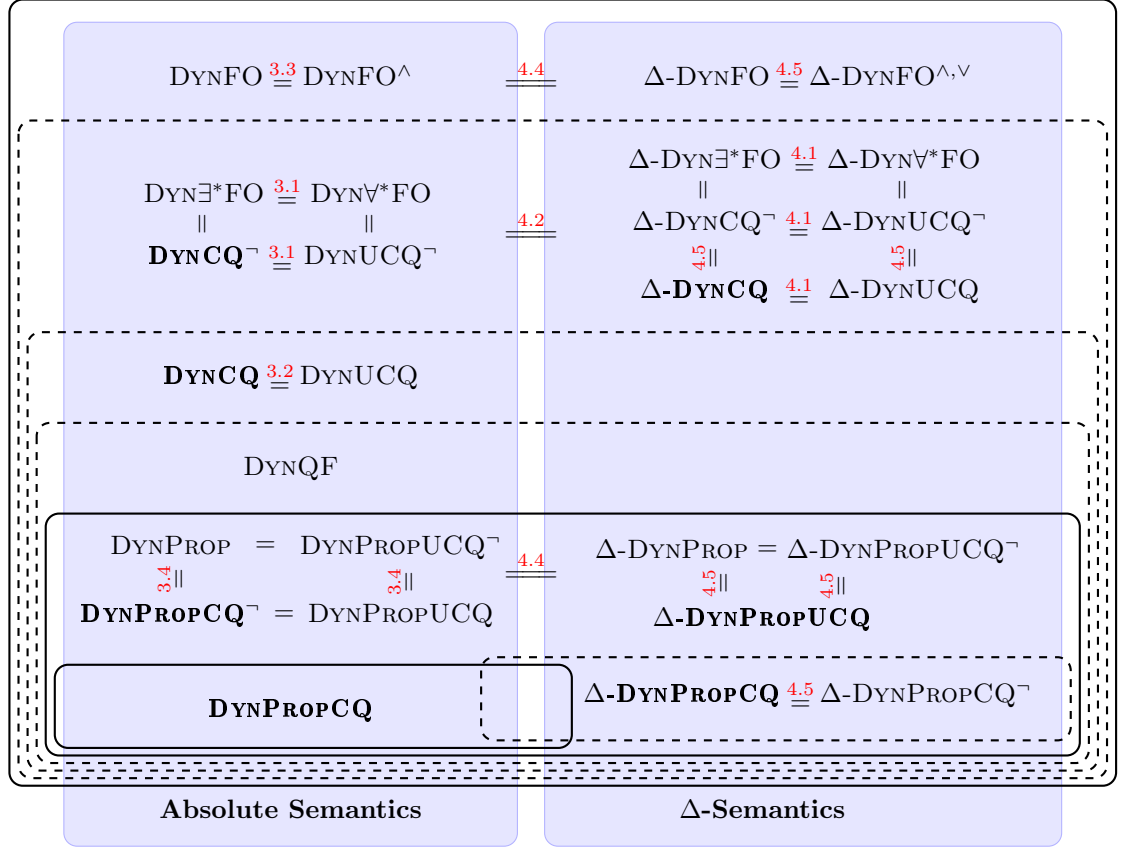


Figure 1: Hierarchy of fragments of DYNFO. Solid lines are strict separations.

**Goal 2.** *Understand the relationship between absolute semantics and  $\Delta$ -semantics for conjunctive queries and their variants.*

In this work we contribute to achieve those two goals as follows.

**Contributions.** For an overview of the relationship of the various dynamic classes of conjunctive queries we refer to Figure 1.

The distinctness of the underlying static query classes does not translate into the dynamic setting:<sup>2</sup>

- We show that, in many cases, the addition of the union-operator does not yield additional expressive power in the dynamic setting, for example,

<sup>2</sup>The notation for classes will be formally introduced in Sections 2 and 4. In general,  $\Delta$  indicates  $\Delta$ -semantics, the absence of  $\Delta$  indicates absolute semantics.

$\text{DYNUCQ}^\neg = \text{DYN}^\neg\text{CQ}^\neg$ ,  $\text{DYNUCQ} = \text{DYN}^\neg\text{CQ}$ , and  $\text{DYNPROPUCQ}^\neg = \text{DYNPROP}^\neg\text{CQ}^\neg$ , where PROP indicates classes without quantifiers.

- Furthermore, negation often does not increase the expressive power of an update language, e.g. we have  $\text{DYNPROPUCQ}^\neg = \text{DYNPROPUCQ}$  and  $\Delta\text{-DYN}^\neg\text{CQ}^\neg = \Delta\text{-DYN}^\neg\text{CQ}$ .
- Finally, often quantifiers can be replaced by their dual quantifiers, e.g.  $\text{DYN}\exists^*\text{FO}(= \text{DYNUCQ}^\neg) = \text{DYN}\forall^*\text{FO}$ .

Whether  $\text{DYN}^\neg\text{CQ}^\neg = \text{DYN}^\neg\text{CQ}$  remains open. However, first-steps towards the separation of the remaining fragments have been taken. Already in [21] the authors proved that dynamic conjunctive queries without negations and quantifiers are strictly weaker than the quantifier-free fragment of  $\text{DYNFO}$ . Here we continue this work:

- By showing that  $\text{DYNQF}$ , the extension of  $\text{DYNPROP}$  that allows auxiliary functions, is contained in  $\text{DYN}^\neg\text{CQ}$ , we can separate the classes  $\text{DYNPROP}$  and  $\text{DYN}^\neg\text{CQ}$ .

Furthermore, we show that dynamic conjunctive queries extended by negations capture all first-order queries:

- We characterize the class of first-order queries as the class maintainable by non-recursive dynamic  $\text{DYN}\exists^*\text{FO}$ -programs with a single existential quantifier per update formula. This implies that dynamic conjunctive queries extended by negations can maintain all first-order queries.

For the second goal, the main finding is that the difference between absolute and  $\Delta$ -semantics is much smaller than we had expected.

- The dynamic classes corresponding to  $\text{FO}$ ,  $\text{CQ}^\neg$  and  $\text{PROP}$  yield the same expressive power with respect to absolute and  $\Delta$ -semantics.
- It turns out that conjunctive queries and conjunctive queries with negation coincide with respect to  $\Delta$ -semantics, that is, in particular,  $\Delta\text{-DYN}^\neg\text{CQ} = \Delta\text{-DYN}^\neg\text{CQ}^\neg$  and thus, also  $\Delta\text{-DYN}^\neg\text{CQ} = \text{DYNUCQ}^\neg$ .

**Choice of setting.** The concrete settings under which dynamic complexity has been studied in the literature slightly differ in several aspects. We shortly discuss the most important aspects, what choice we took for this work and why we made this choice.

An important aspect is whether to use a finite and fixed domain, an active domain or an infinite domain. In this work, we follow the framework of Patnaik and Immerman in which the domain is finite and fixed [16]. To maintain a query, a dynamic program has to work uniformly for all domains. This fixed domain framework for a dynamic setting might appear counterintuitive at first sight. However, it allows to study the underlying dynamic mechanisms of dynamic programs, in particular when one is interested to develop lower bound methods.

Fixed domains also offer a strong connection to logics and circuit complexity. In incremental evaluation systems (IES), a framework proposed by Dong and Topor [7], active domains are used. First-order incremental evaluation systems (FOIES), introduced by Dong and Su in subsequent work [4], have a very close connection to DYNFO. This setting is a little closer to real database systems but most results in dynamic complexity hold equally in both frameworks.

Another parameter to choose is how the auxiliary data is initialized. In the setting of Patnaik and Immerman, dynamic programs start from empty databases and the auxiliary data is either initialized by a polynomial time computation or by a formula from the same class as the update formulas. Later this was generalized by Weber and the second author by proposing that dynamic programs start from an arbitrary initial database and auxiliary data initialized by a mapping computable in some given complexity class [18].

In the present work, we allow for arbitrary initialization mappings. This is motivated by our long term goal to develop lower bound techniques for dynamic programs. While lower bounds in settings with restricted initialization might depend on this restriction, an inexpressibility result in the setting with arbitrary initialization, on the other hand, really shows that a query cannot be *maintained*. A result like  $\text{DYNUCQ} = \text{DYNCQ}$  is helpful for the development of lower bound techniques, as it shows that for proving lower bounds for DYNUCQ it is sufficient to consider DYNCQ programs — but also that one has to be aware that lower bounds for DYNCQ are as hard as lower bounds for DYNUCQ. However, though all our results are stated for arbitrary initialization mappings, they also hold in the setting with empty initial database and first-order initialization for the auxiliary data. On the other hand, some proofs do not carry over to the strict setting of Patnaik and Immerman where, in a dynamic class  $\text{DYNC}$ , only  $\mathcal{C}$  initializations are allowed.

**Related work.** We next discuss some further related work, beyond what we already mentioned above. The expressivity of first-order logic in the dynamic complexity frameworks discussed above has been studied a lot (see e.g. [16, 6, 8, 12, 13, 18, 10, 9]). Most results focus on showing that a problem from some static complexity class can be dynamically maintained by programs of a weaker query class. Some lower bounds have been achieved as well (see e.g. [2, 3, 6, 9, 10, 21]). Many other aspects such as the arity of auxiliary relations (see e.g. [6, 13]), whether the auxiliary relations are determined by the current structure (see e.g. [16, 5, 10]), and the presence of an order (see e.g. [10]) have been studied.

An algebraic perspective of incremental view maintenance under  $\Delta$ -semantics has been studied in [14]. Parts of the latter work have also been implemented, see e.g. [15].

**Outline.** In Section 2 we define our dynamic setting more precisely. In Section 3 the concrete dynamic complexity classes under consideration are introduced. Results for the collapse of classes are obtained in Section 3.1 and results for the separation classes are proved in Section 3.2. The alternative

$\Delta$ -semantics is introduced and studied in Section 4. In Section 5 we give the dynamic characterization of first-order logic. We conclude with a discussion and a first step toward separations in Section 6.

**Acknowledgement.** We thank Nils Vortmeier for careful proofreading. Further we are grateful to the anonymous reviewers of this as well as preceeding work for several very helpful comments.

## 2 Dynamic setting

In this section, we introduce the basic concepts and fix our notation. We mainly borrow it from our previous work [21].

A *dynamic instance* of a query  $Q$  is a pair  $(\mathcal{D}, \alpha)$ , where  $\mathcal{D}$  is a database over a finite domain  $D$  and  $\alpha$  is a sequence of modifications to  $\mathcal{D}$ , i.e. a sequence of insertions and deletions of tuples over  $D$ . The dynamic query  $\text{DYN}(Q)$  yields as result the relation that is obtained by first applying the modifications from  $\alpha$  to  $\mathcal{D}$  and then evaluating the query  $Q$  on the resulting database.

The database resulting from applying a modification  $\delta$  to a database  $\mathcal{D}$  is denoted by  $\delta(\mathcal{D})$ . The result  $\alpha(\mathcal{D})$  of applying a sequence of modifications  $\alpha = \delta_1 \dots \delta_m$  to a database  $\mathcal{D}$  is defined by  $\alpha(\mathcal{D}) \stackrel{\text{def}}{=} \delta_m(\dots(\delta_1(\mathcal{D}))\dots)$ .

Dynamic programs, to be defined next, consist of an initialization mechanism and an update program. The former yields, for every (input) database  $\mathcal{D}$ , an initial state with initial auxiliary data. The latter defines the new state of the dynamic program for each possible modification  $\delta$ .

A *dynamic schema* is a tuple<sup>3</sup>  $(\tau_{\text{in}}, \tau_{\text{aux}})$  where  $\tau_{\text{in}}$  and  $\tau_{\text{aux}}$  are the schemas of the input database and the auxiliary database, respectively. In this work all schemata are purely relational, although all results also hold for input schemas with constants. We always let  $\tau \stackrel{\text{def}}{=} \tau_{\text{in}} \cup \tau_{\text{aux}}$ .

**Definition 1.** (Update program) An *update program*  $\mathcal{P}$  over dynamic schema  $(\tau_{\text{in}}, \tau_{\text{aux}})$  is a set of first-order formulas (called *update formulas* in the following) that contains, for every  $R \in \tau_{\text{aux}}$  and every  $\delta \in \{\text{INS}_S, \text{DEL}_S\}$  with  $S \in \tau_{\text{in}}$ , an update formula  $\phi_\delta^R(\vec{x}; \vec{y})$  over the schema  $\tau$  where  $\vec{x}$  and  $\vec{y}$  have the same arity as  $S$  and  $R$ , respectively.

A *program state*  $\mathcal{S}$  over dynamic schema  $(\tau_{\text{in}}, \tau_{\text{aux}})$  is a structure  $(D, \mathcal{I}, \mathcal{A})$  where  $D$  is a finite domain,  $\mathcal{I}$  is a database over the input schema (the *current database*) and  $\mathcal{A}$  is a database over the auxiliary schema (the *auxiliary database*). The *semantics of update programs* is as follows. For a modification  $\delta(\vec{a})$ , where  $\vec{a}$  is a tuple over  $D$ , and program state  $\mathcal{S} = (D, \mathcal{I}, \mathcal{A})$  we denote by  $P_\delta(\mathcal{S})$  the state  $(D, \delta(\mathcal{I}), \mathcal{A}')$ , where  $\mathcal{A}'$  consists of relations  $R' \stackrel{\text{def}}{=} \{\vec{b} \mid \mathcal{S} \models \phi_\delta^R(\vec{a}; \vec{b})\}$ . The effect  $P_\alpha(\mathcal{S})$  of a modification sequence  $\alpha = \delta_1 \dots \delta_m$  to a state  $\mathcal{S}$  is the state  $P_{\delta_m}(\dots(P_{\delta_1}(\mathcal{S}))\dots)$ .

---

<sup>3</sup>In [21] a dynamic schema had an additional schema for an extra database with built-in relations. As here we do not restrict auxiliary relations in any way and allow arbitrary initialization, we do not need built-in relations.

**Definition 2.** (Dynamic program) A *dynamic program* is a triple  $(P, \text{INIT}, Q)$ , where

- $P$  is an update program over some dynamic schema  $(\tau_{\text{in}}, \tau_{\text{aux}})$ ,
- $\text{INIT}$  is a mapping that maps  $\tau_{\text{in}}$ -databases to  $\tau_{\text{aux}}$ -databases, and
- $Q \in \tau_{\text{aux}}$  is a designated *query symbol*.

A dynamic program  $\mathcal{P} = (P, \text{INIT}, Q)$  *maintains* a dynamic query  $\text{DYN}(Q)$  if, for every dynamic instance  $(\mathcal{D}, \alpha)$ , the relation  $Q(\alpha(\mathcal{D}))$  coincides with the query relation  $Q^{\mathcal{S}}$  in the state  $\mathcal{S} = P_{\alpha}(\mathcal{S}_{\text{INIT}}(\mathcal{D}))$ , where  $\mathcal{S}_{\text{INIT}}(\mathcal{D})$  is the initial state, i.e.  $\mathcal{S}_{\text{INIT}}(\mathcal{D}) \stackrel{\text{def}}{=} (D, \mathcal{D}, \text{INIT}_{\text{aux}}(\mathcal{D}))$ .

Several dynamic settings and restrictions of dynamic programs have been studied in the literature (see e.g. [16, 8, 10, 9]). Possible parameters are, for instance

- the logic in which update formulas are expressed;
- whether, in dynamic instances  $(\mathcal{D}, \alpha)$ , the initial database  $\mathcal{D}$  is always empty; and
- whether the initialization mapping  $\text{INIT}_{\text{aux}}$  is *permutation-invariant* (short: *invariant*), that is, whether  $\pi(\text{INIT}_{\text{aux}}(\mathcal{D})) = \text{INIT}_{\text{aux}}(\pi(\mathcal{D}))$  holds, for every database  $\mathcal{D}$  and permutation  $\pi$  of the domain.

We refer to the introduction for a discussion of the choices made in the following definition.

**Definition 3.** ( $\text{DYN}\mathcal{C}$ ) For a class  $\mathcal{C}$  of formulas, let  $\text{DYN}\mathcal{C}$  be the class of all dynamic queries that can be maintained by dynamic programs with formulas from  $\mathcal{C}$  and arbitrary initialization mapping.

In particular  $\text{DYNFO}$  is the class of all dynamic queries that can be maintained by first-order update formulas.  $\text{DYNPROP}$  is the subclass of  $\text{DYNFO}$ , where update formulas are not allowed to use quantifiers.

We note that arbitrary, (possibly) non-uniform initialization mappings permit to maintain undecidable queries, even when the logic for expressing update formulas is very weak.

Allowing arbitrary initialization mappings in Definition 7 helps us to concentrate on the *maintenance* aspect of dynamic complexity and it helps keeping proofs short. All our results also hold for FO-definable initialization mappings on ordered domains.

**Example 1.** The transitive closure of an acyclic graph can be maintained in  $\text{DYNFO}$ . We follow the argument from [16] and construct a dynamic  $\text{DYNFO}$ -program with one binary auxiliary relation  $T$  which is intended to store the transitive closure.

Insertions can be handled straightforwardly. After inserting an edge  $(u, v)$  there is a path from  $x$  to  $y$  if, before the insertion, there has been a path from  $x$  to  $y$  or there have been paths from  $x$  to  $u$  and from  $v$  to  $y$ . There is a path  $p$  from  $x$  to  $y$  after deleting an edge  $(u, v)$  if there was a path from  $x$  to  $y$  before the deletion (1) but there was no such path via  $(u, v)$ , or (2) there is an edge  $(z, z')$  on  $p$  such that  $u$  can be reached from  $z$  but not from  $z'$ . If there is still a



path  $p$  from  $x$  to  $y$ , such an edge  $(z, z')$  must exist on  $p$ , as otherwise  $u$  would be reachable from  $y$  contradicting acyclicity. All conditions can be checked using the transitive closure of the graph before the deletion of  $(u, v)$ . The update formulas for  $T$  are as follows:

$$\begin{aligned}\phi_{\text{INSE}}^T(u, v; x, y) &\stackrel{\text{def}}{=} T(x, y) \vee (T(x, u) \wedge T(v, y)) \\ \phi_{\text{DELE}}^T(u, v; x, y) &\stackrel{\text{def}}{=} T(x, y) \wedge \left( (\neg T(x, u) \vee \neg T(v, y)) \right. \\ &\quad \vee \exists z \exists z' (T(x, z) \wedge E(z, z') \wedge (z \neq u \vee z' \neq v) \\ &\quad \left. \wedge T(z', y) \wedge T(z, u) \wedge \neg T(z', u)) \right)\end{aligned}$$

□

The following notion will be useful at several occasions. The *dependency graph* of a dynamic program  $\mathcal{P}$  with auxiliary schema  $\tau$  has vertex set  $V = \tau$  and an edge  $(R, R')$  if the relation symbol  $R'$  occurs in one of the update formulas for  $R$ . The *deletion dependency graph* is defined as the dependency graph except that only update formulas for delete operations are taken into account.

### 3 Dynamic conjunctive queries

In this section we study the relationship between the variants of dynamic conjunctive queries.

We first give formal definitions of the classes of queries we are interested in:

- CQ: the class of conjunctive queries, that is, queries expressible by first-order formulas of the form  $\varphi(\vec{x}) = \exists \vec{y} \psi$ , where  $\psi$  is a conjunction of atomic formulas.
- UCQ: the class of all unions of conjunctive queries, that is, queries expressible by formulas of the form  $\bigvee_i \exists \vec{x} \psi_i$ , where each  $\psi_i$  is a conjunction of atomic formulas.

We note that safety of queries is not an issue in this paper: we use queries as update formulas only and we can always assume that, for each required arity, there is an auxiliary “universal” relation containing all tuples of this arity over the active domain which could be used to make queries syntactically safe.

The classes CQ and UCQ can be extended by additionally allowing negated atoms, resulting in  $\text{CQ}^-$  and  $\text{UCQ}^-$ ; or they can be restricted by disallowing quantification, resulting in  $\text{PROPCQ}$ ,  $\text{PROPUCQ}$ ,  $\text{PROPCQ}^-$  and  $\text{PROPUCQ}^-$ . It is well known that  $\text{UCQ}^-$  and  $\exists^* \text{FO}$ , the class of queries expressible by existential first-order formulas, coincide, but otherwise, all these classes are distinct. Furthermore, other quantification patterns than  $\exists^*$  can be considered, like  $\forall^*$  or arbitrary quantification.

The dynamic program maintaining the transitive closure for acyclic graphs in Example 1 is actually a  $\text{DYNUCQ}^-$ -program.

The main goal of this section is to show that the relationship of all these classes in the dynamic setting is much simpler than in the static setting. In the first part of the section we prove that many dynamic classes collapse, as indicated in the left part of Figure 1. In the second part we show that the dynamic classes DYNPROPCQ, DYNPROP and DYNCQ can be separated. Whether DYNCQ, DYNUCQ and DYNFO can be separated remains open.

### 3.1 Collapse results

In this section we prove that dynamic classes collapse as indicated in the left part of Figure 1. More precisely, we show the following theorems. The main results of this section are the following two theorems regarding the second and the third fragment in the left part of Figure 1.

**Theorem 3.1.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in DYNUCQ $^\neg$ .
- (b)  $\mathcal{Q}$  can be maintained in DYNCQ $^\neg$ .
- (c)  $\mathcal{Q}$  can be maintained in DYN $\exists^*$ FO.
- (d)  $\mathcal{Q}$  can be maintained in DYN $\forall^*$ FO.

**Theorem 3.2.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in DYNUCQ.
- (b)  $\mathcal{Q}$  can be maintained in DYNCQ.

Using the same technique as is used for removing unions from dynamic unions of conjunctive queries, a normal form for DYNFO can be obtained. The class DYNFO $^\wedge$  contains all queries maintainable by a program whose update formulas are in prenex normal form where the quantifier-free part is a conjunction of atoms.

**Theorem 3.3.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in DYNFO.
- (b)  $\mathcal{Q}$  can be maintained in DYNFO $^\wedge$ .

Further we prove the following result for the quantifier-free variants of dynamic conjunctive queries.

**Theorem 3.4.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in DYNPROP.
- (b)  $\mathcal{Q}$  can be maintained in DYNPROPUCQ $^\neg$ .
- (c)  $\mathcal{Q}$  can be maintained in DYNPROPCQ $^\neg$ .

(d)  $\mathcal{Q}$  can be maintained in DYNUCQ.

Before we turn to the proofs of the theorems, we discuss the proof techniques that will be used.

For showing that a class  $\text{DYN}\mathcal{C}$  of queries is contained in a class  $\text{DYN}\mathcal{C}'$ , it is sufficient to construct, for every dynamic program with update queries from class  $\mathcal{C}$ , an equivalent dynamic program with update queries from class  $\mathcal{C}'$ . In cases where  $\mathcal{C}' \subset \mathcal{C}$  this can also be seen as constructing a  $\mathcal{C}'$ -normal form for  $\mathcal{C}$ -programs.

Most of the proofs for the collapse of two dynamic classes in this paper are not very deep. Indeed, most of them use one or more of the following three (easy) techniques.

The *replacement technique* ([21]) is used to remove subformulas of a certain kind from update formulas and to replace their “meaning” by additional auxiliary relations. In this way, we often can remove negations (choose negative literals as subformulas, see the proof of Lemma 3.6) and disjunctions (see proof of Lemma 3.8) from update formulas.

The *preprocessing technique* is used to convert (more) complicated update formulas into easier update formulas by splitting the computation performed by the complicated update formula into two parts; one of them performed by the initialization mapping and stored in an additional auxiliary relation, the other one performed by the easier update formula using the pre-computed auxiliary relation. Applications of this technique are the removal of unions from dynamic unions of conjunctive queries (see example below) as well as proving the equivalence of semantics for dynamic conjunctive queries with negations (see Lemma 4.6).

**Example 2.** We consider the update formula

$$\phi_\delta^R(u; x) = \exists y (U(x, y) \vee V(x, u))$$

for a unary relational symbol  $R$ . We aim at an equivalent update formula  $\psi_\delta^R(u; x)$  without disjunction. The idea is to store a ‘disjunction blue print’ in a precomputed auxiliary relation  $T$  and to use existential quantification to guess which disjunct becomes true.

In this example, we assume that in every state of the dynamic program on every database, both the interpretations of  $U$  and  $V$  are always non-empty sets.<sup>4</sup> Then,  $\phi_\delta^R(u; x)$  can be replaced by

$$\exists y \exists z_1 \exists z_2 \exists z_3 \exists z_4 (U(z_1, z_2) \wedge V(z_3, z_4) \wedge T(z_1, z_2, z_3, z_4, x, y, u))$$

where  $T$  is an additional auxiliary relation symbol which is interpreted, in every state  $\mathcal{S}$ , by a 7-ary relation  $T^\mathcal{S}$  containing all tuples  $(a_1, \dots, a_7)$  with  $(a_1, a_2) = (a_5, a_6)$  or  $(a_3, a_4) = (a_5, a_7)$ . Thus  $T^\mathcal{S}$  ensures that either the values chosen for  $z_1, z_2$  coincide with the values of  $x, y$  or the values of  $z_3, z_4$  coincide with  $x, u$ .

---

<sup>4</sup>This assumption will eventually be removed in the proof.

Therefore, the initialization mapping initializes  $T$  with the result of the query

$$\mathcal{Q}_T(z_1, z_2, z_3, z_4, x, y, u) \stackrel{\text{def}}{=} ((z_1, z_2) = (x, y) \vee (z_3, z_4) = (x, u)).$$

Observe that this approach fails when  $U$  or  $V$  are interpreted by empty relations. In order to cover empty relations as well, some extra work needs to be done (see proof of Lemma 3.8).

The *squirrel technique* maintains additional auxiliary relations that reflect the state of some auxiliary relation after every possible single modification (or short modification sequence).<sup>5</sup> For example, if a dynamic program contains a relation symbol  $R$  then a fresh relation symbol  $R_{\text{INS}}$  can be used, such that the interpretation of  $R_{\text{INS}}$  contains the content of  $R$  after modification INS (for every possible insertion tuple). Of course,  $R_{\text{INS}}$  has higher arity than  $R$ , as it takes the actual inserted tuple into account. Sample applications of this technique are the removal of quantifiers from *some* update formulas (see the following example and Lemma 3.5) and the maintenance of first-order queries in  $\text{DYN}CQ^-$  (see Theorem 5.1).

**Example 3.** Consider the update formula

$$\phi_{\text{INS}}^Q(u_1; x) = \exists y(Q(x) \vee \neg S(u_1, y))$$

for the query symbol  $Q$  of some dynamic program  $\mathcal{P}$ . In order to obtain a quantifier-free update formula for  $Q$  after insertion of an arbitrary tuple we maintain the relation  $Q_{\text{INS}}(\cdot, \cdot)$  that contains a tuple  $(a, b)$  if and only if  $b$  would be in  $Q$  in the next state, after insertion of  $a$ . Similarly for  $S$  and deletions.

Then the update formula  $\phi_{\text{INS}}^Q$  can be replaced by the quantifier-free formula  $\phi_{\text{INS}}^Q(u_1; x) \stackrel{\text{def}}{=} Q_{\text{INS}}(u_1, x)$ . The relation  $Q_{\text{INS}}$  can be updated via

$$\begin{aligned} \phi_{\text{INS}}^{Q_{\text{INS}}}(u_0; u_1, x) &\stackrel{\text{def}}{=} \exists y(Q_{\text{INS}}(u_0, x) \vee \neg S_{\text{INS}}(u_0, u_1, y)) \\ \phi_{\text{DEL}}^{Q_{\text{INS}}}(u_0; u_1, x) &\stackrel{\text{def}}{=} \exists y(Q_{\text{DEL}}(u_0, x) \vee \neg S_{\text{DEL}}(u_0, u_1, y)) \end{aligned}$$

and similarly, for the other new auxiliary relations.

We note that, in this example, the application of the technique does not eliminate all quantifiers in the program (in fact, it removes one and introduces two new formulas with quantifiers), but it removes quantification from the update formula for a *particular relation*. Removing quantification from the update formulas of the query relation will turn out to be useful in the proofs of Lemmata 3.8, 3.10 and 4.6.

This concludes the description of the techniques. In the following, as a preparatory step, we show how to remove quantification from the update formulas of the query relation. Afterwards we exhibit constructions for removing negations and unions from some dynamic fragments. Finally we present a construction for switching quantifiers in dynamic programs, that is, e.g., for constructing an  $\exists^*\text{FO}$ -program from an  $\forall^*\text{FO}$ -program.

---

<sup>5</sup>Squirrels usually make provisions for every possible future.

The following technical lemma shows how to remove quantifiers from the update formulas of the query relation. For an arbitrary quantifier prefix  $\mathbb{Q} \in \{\exists, \forall\}^*$  let  $\mathbb{Q}\text{FO}$  be the class of queries expressible by formulas with quantifier prefix  $\mathbb{Q}$ . If  $\mathbb{Q}$  is a substring of  $\mathbb{Q}'$  and  $Q$  is a query in  $\mathbb{Q}\text{FO}$  then trivially  $Q$  is in  $\mathbb{Q}'\text{FO}$  as well.

**Lemma 3.5.** *Let  $\mathbb{Q}$  be an arbitrary quantifier prefix. For every  $\text{DYN}\mathbb{Q}\text{FO}$ -program there is an equivalent  $\text{DYN}\mathbb{Q}\text{FO}$ -program  $\mathcal{P}$  such that the update formulas for the designated query symbol of  $\mathcal{P}$  consist of a single atom.*

*Proof.* We follow the approach from Example 3. For ease of presentation we fix the input schema to be  $\tau_{\text{in}} = \{E\}$  where  $E$  is a binary relation symbol; the proof can be easily adapted to arbitrary input schemas.

Let  $\mathcal{P}$  be a  $\text{DYN}\mathcal{C}$ -program over auxiliary schema  $\tau$  with designated query symbol  $Q$ . We construct an equivalent  $\text{DYN}\mathcal{C}$  program  $\mathcal{P}'$  over schema  $\tau'$  where  $\tau'$  contains a designated query symbol  $Q'$  and a  $(k+2)$ -ary relation symbol  $R_\delta$  for every  $k$ -ary  $R \in \tau$  and every  $\delta \in \{\text{INS}, \text{DEL}\}$ .

The idea is that  $R_\delta$  shall reflect the content of  $R$  in the next state, for each possible modification of the kind  $\delta$ . More precisely, let  $G = (E, V)$  be a graph,  $\alpha$  a sequence of modifications,  $\beta = \delta(\vec{e})$  a modification with  $\delta \in \{\text{INS}, \text{DEL}\}$  and  $\vec{e} \in V^2$ . If  $\mathcal{S}$  is the state obtained by  $\mathcal{P}$  after applying  $\alpha\beta$  to  $G$ , i.e.  $\mathcal{S} = \mathcal{P}_{\alpha\beta}(\text{INIT}(G))$ , and  $\mathcal{S}'$  is the state obtained by  $\mathcal{P}'$  after applying  $\alpha$  to  $G$ , i.e.  $\mathcal{S}' = \mathcal{P}'_\alpha(\text{INIT}'(G))$ , then

$$\vec{a} \in R^{\mathcal{S}} \text{ if and only if } (\vec{e}, \vec{a}) \in R_\delta^{\mathcal{S}'} . \quad (1)$$

Thus for every  $\delta(\vec{e})$  the relation  $R_\delta(\vec{e}, \cdot)$  stores  $R(\cdot)$  after application of  $\delta(\vec{e})$ .

Then the update formula for  $Q'$  after a modification  $\delta$  can be written as follows:

$$\phi_\delta^{Q'}(\vec{u}; \vec{x}) \stackrel{\text{def}}{=} R_\delta(\vec{u}, \vec{x})$$

It remains to explain how to update the relations  $R_\delta$ . We use formulas  $\phi_{\text{INS}}^E$  and  $\phi_{\text{DEL}}^E$  that express the impact of a modification to  $E$ , for example,  $\phi_{\text{INS}}^E(a, b; x, y) = E(x, y) \vee (a = x \wedge b = y)$ . By  $\phi_{\delta_1}^R[\tau \rightarrow \tau_{\delta_0}](\vec{u}_0; \vec{u}_1, \vec{x})$  we denote the formula obtained from  $\phi_{\delta_1}^R(\vec{u}_1; \vec{x})$  by replacing every atom  $S(\vec{z})$  with  $S \in \tau$  by  $S_{\delta_0}(\vec{u}_0, \vec{z})$ . Then the update formula for  $R_{\delta_1}$  is

$$\phi_{\delta_0}^{R_{\delta_1}}(\vec{u}_0; \vec{u}_1, \vec{x}) \stackrel{\text{def}}{=} \phi_{\delta_1}^R[\tau \rightarrow \tau_{\delta_0}](\vec{u}_0; \vec{u}_1, \vec{x}).$$

The initialization mapping of  $\mathcal{P}'$  is as follows. The query symbol  $Q'$  is initialized like  $Q$  in  $\mathcal{P}$ . For every graph  $G$  the relation symbol  $R_\delta \in \tau'$  is initialized as

$$\bigcup_{\vec{e} \in V^2} \{\vec{e}\} \times \mathcal{P}_{\delta(\vec{e})}(\text{INIT}(G)) \upharpoonright R_\delta$$

where  $\mathcal{P}_{\delta(\vec{e})}(\text{INIT}(G)) \upharpoonright R_\delta$  denotes the relation  $R_\delta$  in state  $\mathcal{P}_{\delta(\vec{e})}(\text{INIT}(G))$ .

The correctness of this construction is proved inductively over the length of modification sequences by showing that states of  $\mathcal{P}'$  *simulate* states of  $\mathcal{P}$  as specified by (1).

Therefore, let  $G$  be a graph and  $\alpha = \alpha_1 \dots \alpha_i$  a modification sequence with  $\alpha_i = \delta_i(\vec{e}_i)$ . Further let  $\mathcal{S}_j$  and  $\mathcal{S}'_j$  be the states obtained by  $\mathcal{P}$  and  $\mathcal{P}'$ , respectively, after application of  $\alpha_1 \dots \alpha_i$ .

If  $\alpha$  is of length 0 and  $\beta$  is an arbitrary modification with  $\beta = \delta(\vec{e})$  then  $\mathcal{S} \stackrel{\text{def}}{=} \mathcal{P}_\beta(\mathcal{S}_0)$  and  $\mathcal{S}' \stackrel{\text{def}}{=} \mathcal{S}'_0$  satisfy (1) thanks to the definition of the initialization mapping of  $\mathcal{P}'$ .

If  $\alpha$  is of length  $i \geq 1$  then, by induction hypothesis, the states  $\mathcal{S} \stackrel{\text{def}}{=} \mathcal{S}_i = \mathcal{P}_{\alpha_i}(\mathcal{S}_{i-1})$  and  $\mathcal{S}' \stackrel{\text{def}}{=} \mathcal{S}'_{i-1}$  satisfy (1) that is

$$\vec{a} \in R^{\mathcal{S}} \text{ if and only if } (\vec{e}_i, \vec{a}) \in R_{\delta_i}^{\mathcal{S}'} \quad (2)$$

for all relations  $R$  and  $R_{\delta_i}$ .

Now, let  $\beta = \delta(\vec{e})$  be an arbitrary modification. Further let  $\mathcal{T} \stackrel{\text{def}}{=} \mathcal{P}_\beta(\mathcal{S})$  and  $\mathcal{T}' \stackrel{\text{def}}{=} \mathcal{P}'_{\alpha_i}(\mathcal{S}')$ . By definition,  $\vec{b} \in R^{\mathcal{T}}$  if and only if

$$(R^{\mathcal{S}}, \{\vec{u}_1 \mapsto \vec{e}, \vec{x} \mapsto \vec{b}\}) \models \phi_{\delta}^R(\vec{u}_1; \vec{x}).$$

Thanks to (2) and the definition of  $\phi_{\delta_0}^{R_{\delta_1}}$  this is equivalent to

$$(R^{\mathcal{S}'}, \{\vec{u}_0 \mapsto \vec{e}_i, \vec{u}_1 \mapsto \vec{e}, \vec{x} \mapsto \vec{b}\}) \models \phi_{\delta_i}^R[\tau \rightarrow \tau_{\delta_0}](\vec{u}_0; \vec{u}_1, \vec{x}).$$

By definition this is equivalent to  $(\vec{e}, \vec{b}) \in R^{\mathcal{T}'}$ .  $\square$

Now we turn towards constructions for removing negations and unions from certain fragments. We start by exhibiting negation-free normal forms for DYNFO and for DYNPROP.

**Lemma 3.6.** (a) *Every DYNFO-program has an equivalent negation-free DYNFO-program.*

(b) *Every DYNPROP-program has an equivalent DYNPROPUCQ-program.*

*Proof.* This theorem is a generalization of Theorem 6.6 from [13]. Given a dynamic program  $\mathcal{P}$ , the simple idea is to maintain, for every auxiliary relation  $R$  of  $\mathcal{P}$ , an additional auxiliary relation  $\hat{R}$  for the complement of  $R$ .

We make this more precise. In the following we prove (a). As the construction does not introduce quantifiers, it can be used for (b) as well.

Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a DYNFO-program over schema  $\tau$ . We assume, without loss of generality, that  $\mathcal{P}$  is in negation normal form. Further we assume, for ease of presentation, that the input relations have update formulas as well<sup>6</sup>.

We construct a negation-free DYNFO-program equivalent to  $\mathcal{P}$  that uses the schema  $\tau \cup \hat{\tau} \cup \{\hat{=}\}$  where  $\hat{\tau}$  contains for every relation symbol  $R \in \tau$  a fresh relation symbol  $\hat{R}$  of equal arity. Recall that  $\tau$  includes the input schema and the auxiliary schema. The idea is to maintain in  $\hat{R}^{\mathcal{S}}$  the negation of  $R^{\mathcal{S}}$ , for all states  $\mathcal{S}$ . Further  $\hat{=}$  shall always contain the complement of  $=$ .

<sup>6</sup>E.g. if the input database is a graph, then  $\phi_{\text{INS}}^E(a, b; x, y) = E(x, y) \vee (a = x \wedge b = y)$  etc.

In a first step we construct a DYNFO-program  $\mathcal{P}' = (P', \text{INIT}', Q)$  in negation normal form over  $\tau \cup \widehat{\tau} \cup \{\hat{=}\}$  that maintains  $R^S$  and  $\widehat{R}^S$  (but still uses negations). The update formulas for relation symbols  $R \in \tau$  are as in  $\mathcal{P}$ . For every  $\widehat{R} \in \widehat{\tau}$  and every modification  $\delta$ , the update formula  $\phi_\delta^{\widehat{R}}(\vec{x}; \vec{y})$  is the negation normal form<sup>7</sup> of  $\neg\phi_\delta^R(\vec{x}; \vec{y})$ . The relation  $\hat{=}$  never changes. The initialization mapping  $\text{INIT}'$  initializes  $\widehat{R}$  with the complement of  $\text{INIT}(R)$ .

From  $\mathcal{P}'$  we construct a negation-free DYNFO-program  $\mathcal{P}'' = (P'', \text{INIT}'', Q)$ . An update formula  $\phi_\delta^R(\vec{x}; \vec{y})$  for  $\mathcal{P}''$  is obtained from the update formula  $\phi_\delta^R(\vec{x}; \vec{y})$  for  $\mathcal{P}'$  by replacing all negative literals  $\neg S$  by  $\widehat{S}$ . The initialization mapping of  $\mathcal{P}''$  is the same as for  $\mathcal{P}'$ .

The equivalence of  $\mathcal{P}$  and  $\mathcal{P}''$  can be proved by an induction over the length of modification sequences. □

Two different techniques are used for removing unions. We start by giving a disjunction-free normal form for DYNPROP.

**Lemma 3.7.** *Every DYNPROP-program has an equivalent DYNPROPCQ<sup>-</sup>-program.*

*Proof.* Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a DYNPROP-program over schema  $\tau$ . We assume, without loss of generality, that  $\tau$  contains, for every relation symbol  $R$ , a relation symbol  $\widehat{R}$  and that  $\mathcal{P}$  ensures that  $\widehat{R}^S$  is the complement of  $R^S$  for every state  $S$ . This can be achieved by using the same technique as in Lemma 3.6. Further we assume that all update formulas of  $\mathcal{P}$  are in conjunctive normal form.

The conjunctive DYNPROP-program we are going to construct is over schema  $\tau \cup \tau'$  where  $\tau'$  contains a fresh relation symbol  $R_{-C}$  for every clause  $C$  occurring in some update formula of  $\mathcal{P}$ . The goal of the construction is to ensure that  $R_{-C}^S(\vec{z})$  holds if and only if  $\neg C(\vec{z})$  is true in state  $S$ . Then an update formula  $\phi = C_1(\vec{x}_1) \wedge \dots \wedge C_k(\vec{x}_k)$  with clauses  $C_1(\vec{x}_1), \dots, C_k(\vec{x}_k)$  can be replaced by the conjunctive formula  $\neg R_{-C_1}(\vec{x}_1) \wedge \dots \wedge \neg R_{-C_k}(\vec{x}_k)$ .

In a first step we construct a DYNPROP-program  $\mathcal{P}' = (P', \text{INIT}', Q)$  in conjunctive normal form that maintains the relations  $R_{-C}^S$ . To this end, let  $C$  be a clause with  $k$  variables and let  $\vec{z}$  be the  $k$ -tuple that contains those variables in the order in which they occur in  $C$ . Assume that  $C = L_1(\vec{z}_1) \vee \dots \vee L_l(\vec{z}_l)$  where  $\vec{z}_i \subseteq \vec{z}$  and each  $L_i$  is an atom or a negated atom. Thus  $\neg C \equiv \neg L_1(\vec{z}_1) \wedge \dots \wedge \neg L_l(\vec{z}_l)$ . The relation symbol  $R_{-C}$  is of arity  $k$ . For a modification  $\delta$  the update formula for  $R_{-C}$  is

$$\phi_\delta^{R_{-C}}(\vec{x}; \vec{z}) = \phi_\delta^{X_1}(\vec{x}; \vec{z}_1) \wedge \dots \wedge \phi_\delta^{X_l}(\vec{x}; \vec{z}_l)$$

where  $X_i$  is the relation symbol  $R$  if  $L_i = \neg R$  and  $X_i$  is  $\widehat{R}$  if  $L_i = R$ . Observe that  $\phi_\delta^{R_{-C}}(\vec{x}; \vec{z})$  is in conjunctive normal form, because each  $\phi_\delta^{X_i}(\vec{x}; \vec{z}_i)$  is in conjunctive normal form; further  $\phi_\delta^{R_{-C}}(\vec{x}; \vec{z})$  does not use new clauses.

<sup>7</sup>Observe that this fails for DYNQC<sup>-</sup> and DYNQC<sup>-</sup>.

The initialization mapping  $\text{INIT}'$  extends the initialization mapping  $\text{INIT}$  to the schema  $\tau'$  in a natural way. For a clause  $C$  and input database  $\mathcal{I}$ , a tuple  $\vec{a}$  is in  $\text{INIT}'(R_{-C})$  if and only if  $C$  evaluates to false in  $\text{INIT}(\mathcal{I})$  for  $\vec{a}$ .

The second step is to construct from  $\mathcal{P}'$  the desired conjunctive DYNPROP-program  $\mathcal{P}''$ : every clause  $C$  in every update formula of  $\mathcal{P}'$  is replaced by  $\neg R_{-C}$ . This construction yields a conjunctive program  $\mathcal{P}''$ . The initialization mapping of  $\mathcal{P}''$  is the same as for  $\mathcal{P}'$ .

We sketch the proof that  $\mathcal{P}''$  is equivalent to  $\mathcal{P}$ . The dynamic program  $\mathcal{P}'$  updates relations from  $\tau$  exactly as program  $\mathcal{P}$ . By an induction over the length of modification sequences, one can prove that  $R_{-C}^S(\vec{a})$  holds if and only if  $\neg C(\vec{a})$  is true in state  $\mathcal{S}$  for all tuples  $\vec{a}$ . Thus corresponding update formulas of  $\mathcal{P}$  and  $\mathcal{P}''$  always yield the same result.  $\square$

Now we turn to disjunction-free normalforms for DYNUCQ, DYNUCQ $^\neg$  and negation-free DYNFO. Observe that the idea of the proof of Lemma 3.7 cannot be applied directly since those classes are not closed under negations. Instead the idea is to simulate disjunctions by existential quantifiers.

**Lemma 3.8.** (a) *For every DYNUCQ $^\neg$ -program there is an equivalent DYNCQ $^\neg$ -program.*

(b) *For every DYNUCQ-program there is an equivalent DYNCQ-program.*

(c) *For every DYNFO-program there is an equivalent DYNFO $^\wedge$ -program.*

*Proof.* We first prove the statements for domains with at least two elements and show how to drop this restriction afterwards. The construction uses the idea from Example 2. We give it for (a) but, as it does not introduce any negation operators it works for (b) as well. For (c) it is sufficient to start from a negation-free DYNFO-program by Lemma 3.6; and for those the same construction as for (a) can be used<sup>8</sup>.

Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a DYNUCQ $^\neg$ -program over schema  $\tau$ . Without loss of generality, we assume that the quantifier-free parts of all update formulas of  $\mathcal{P}$  are in disjunctive normal form. We convert  $\mathcal{P}$  into an equivalent DYNCQ $^\neg$ -program  $\mathcal{P}'$  whose update formulas are in prenex normalform with quantifier-free parts of the form  $\bigwedge_i L_i(\vec{x}_i) \wedge T(\vec{y})$ , where  $L_i$  are arbitrary literals over a modified schema  $\hat{\tau}$  and the symbols  $T$  are fresh auxiliary relation symbols. The program  $\mathcal{P}' = (P', \text{INIT}', Q)$  is over schema  $\tau' = \tau \cup \hat{\tau} \cup \tau_T$ , where  $\hat{\tau}$  contains a  $(k+1)$ -ary relation symbols  $\hat{R}_1$  and  $\hat{R}_2$  for every  $k$ -ary relation symbol  $R \in \tau$ ; and  $\tau_T$  contains a relation symbol  $T_{S,\delta}$  for every relation symbol  $S \in \tau \cup \hat{\tau}$  and every modification  $\delta$ .

The intention for relation symbols from  $\tau'$  is as follows. The relation symbols  $R \in \tau$  shall always be interpreted as in  $\mathcal{P}$ . The intention of  $\hat{R}_1 \in \hat{\tau}$  is, on one hand, to contain a “copy of  $R$ ” (in tuples with first component  $c$ , for some fixed

---

<sup>8</sup>More precisely, replace the quantifier prefix  $\exists \vec{y}$  used throughout the construction of (a) by a general quantifier-prefix  $\exists \vec{y}_1 \forall \vec{y}_2 \dots$



element  $c$ ) and on the other hand, to guarantee non-emptiness. The latter is strongly ensured by enforcing all tuples that do *not* have  $c$  as first component to be in  $\widehat{R}$  and by  $|D| \geq 2$ . Similarly  $\widehat{R}_2 \in \widehat{\tau}$  contains a “copy of  $R$ ” but is not the universal relation, that is, it does not contain all tuples. More precisely:

$$\widehat{R}_1^S \stackrel{\text{def}}{=} \{(c, \vec{a}) \mid \vec{a} \in R^S\} \cup \{(d, \vec{a}) \mid d \neq c \text{ and } \vec{a} \in D^k\} \quad (3)$$

$$\widehat{R}_2^S \stackrel{\text{def}}{=} \{(c, \vec{a}) \mid \vec{a} \in R^S\} \quad (4)$$

The relations  $T_{S,\delta}$  will be used as in Example 2.

Now we construct the update formulas for program  $\mathcal{P}'$ . Let  $R \in \tau$  and  $\delta$  be a modification. Further let

$$\phi_\delta^R(\vec{u}; \vec{x}) = \exists \vec{y} (C_1(\vec{u}, \vec{x}, \vec{y}) \vee \dots \vee C_k(\vec{u}, \vec{x}, \vec{y}))$$

be the update formula of  $R$  with respect to  $\delta$  in  $\mathcal{P}$ , where every  $C_i$  is a conjunction of literals. For

$$C_i(\vec{u}, \vec{x}, \vec{y}) = L_1(\vec{v}_1) \wedge \dots \wedge L_m(\vec{v}_l)$$

we define

$$\widehat{C}_i(v, \vec{u}, \vec{x}, \vec{y}) \stackrel{\text{def}}{=} \widehat{L}_1(v, \vec{v}_1) \wedge \dots \wedge \widehat{L}_m(v, \vec{v}_l)$$

where  $\widehat{L}_j = \widehat{R}_1$  if  $L_j = R$  and  $\widehat{L}_j = \neg \widehat{R}_2$  if  $L_j = \neg R$ .

The update formula  $\psi_\delta^{\widehat{R}_i}(\vec{u}; x', \vec{x})$  for  $\widehat{R}_i \in \widehat{\tau}$  in  $\mathcal{P}'$  is

$$\begin{aligned} \psi_\delta^{\widehat{R}_i}(\vec{u}; x', \vec{x}) &\stackrel{\text{def}}{=} \exists \vec{y} \exists z'_1 \exists \vec{z}_1 \dots \exists z'_k \exists \vec{z}_k \\ &\quad \left( \widehat{C}_1(z'_1, \vec{z}_1) \wedge \dots \wedge \widehat{C}_k(z'_k, \vec{z}_k) \right. \\ &\quad \left. \wedge T_{\widehat{R}_i, \delta}(\vec{y}, z'_1, \vec{z}_1, \dots, z'_k, \vec{z}_k, \vec{u}, x', \vec{x}) \right). \end{aligned}$$

The update formula  $\psi_\delta^R(\vec{u}; \vec{x})$  for  $R \in \tau$  in  $\mathcal{P}'$  is

$$\begin{aligned} \psi_\delta^R(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \exists \vec{y} \exists z'_1 \exists \vec{z}_1 \dots \exists z'_k \exists \vec{z}_k \\ &\quad \left( \widehat{C}_1(z'_1, \vec{z}_1) \wedge \dots \wedge \widehat{C}_k(z'_k, \vec{z}_k) \right. \\ &\quad \left. \wedge T_{R, \delta}(\vec{y}, z'_1, \vec{z}_1, \dots, z'_k, \vec{z}_k, \vec{u}, \vec{x}) \right). \end{aligned}$$

To ensure equivalence of this program with the original program, the relations  $T_{S,\delta}$  are defined as follows.

- $T_{R,\delta}$  contains all tuples<sup>9</sup>  $(\vec{y}, z'_1, \vec{z}_1, \dots, z'_k, \vec{z}_k, \vec{u}, \vec{x})$ , for which, for some  $i$ ,  $z'_i = c$  and  $\vec{z}_i = (\vec{u}, \vec{x}, \vec{y})$ .
- $T_{\widehat{R}_i, \delta}$  contains all tuples  $(\vec{y}, z'_1, \vec{z}_1, \dots, z'_k, \vec{z}_k, \vec{u}, x', \vec{x})$ , for which

---

<sup>9</sup>For simplicity, we reuse variable names as element names.

- $x' \neq c$ , or
- for some  $j$ ,  $z'_j = c$  and  $\vec{z}_j = (\vec{u}, \vec{x}, \vec{y})$ .

These are initialized as intended by simple quantifier-free formulas (but with disjunction). Their interpretation is never changed, that is, for every  $T_{S,\delta}$ , both update formulas reproduce the current value of  $T_{S,\delta}$ .

The initialization for relation symbols from  $\tau$  and  $\hat{\tau}$  is straightforward. Auxiliary relation symbols  $R \in \tau$  are initialized as in  $\mathcal{P}$ ; and auxiliary relation symbols  $\hat{R}_1, \hat{R}_2 \in \hat{\tau}$  are initialized by  $\text{INIT}'$  analogously to  $\text{INIT}$  but respecting Equations (3) and (4).

This concludes the proof of (a), (b) and (c) for domains with at least two elements. The restriction on the size of the domains can be dropped as follows. In all three cases the idea is to make a case distinction on the size of the domain in the update formulas of the designated query symbol.

To this end, we first construct a DYNPROPCQ-program  $\mathcal{P}'' = (\mathcal{P}'', \text{INIT}'', Q'')$  over schema  $\tau''$  with  $\tau' \cap \tau'' = \emptyset$  which is equivalent to  $\mathcal{P}$  over databases with domains of size one. Then we construct a program  $\mathcal{P}'''$  equivalent to  $\mathcal{P}$  by combining the programs  $\mathcal{P}'$  and  $\mathcal{P}''$ .

For the construction of  $\mathcal{P}''$  we observe that every relation of a database over a single element domain  $D = \{a\}$  contains either exactly one tuple, namely  $(a, \dots, a)$ , or no tuple at all. Thus every such relation  $R$  corresponds to a 0-ary relation  $R_0$  where  $R_0$  is true if and only if  $(a, \dots, a) \in R$ . Hence, by Lemma 3.9 (see below), there is a DYNPROPCQ-program equivalent to  $\mathcal{P}$  for databases with domains of size one.

To combine  $\mathcal{P}'$  and  $\mathcal{P}''$  we use two different approaches, one for (a) and one for (b) and (c).

First we consider (a). To this end, we can assume, by Lemma 3.5, that the update formulas for the query relations  $Q'$  and  $Q''$  of  $\mathcal{P}'$  and  $\mathcal{P}''$ , respectively, consist of single atoms. We construct an intermediate program  $\tilde{\mathcal{P}} = (\tilde{\mathcal{P}}, \text{INIT}, \tilde{Q})$  over schema  $\tilde{\tau} = \{\tilde{Q}, U\} \cup \tau' \cup \tau''$  where  $U$  is a fresh 0-ary relation symbol. The intention is that interpretations of symbols in  $\tau'$  and  $\tau''$  are as in  $\mathcal{P}'$  and  $\mathcal{P}''$ , respectively, and that  $U$  is interpreted by true if and only if the domain is of size one. The initializations are accordingly.

Thus all update formulas of  $\tilde{\mathcal{P}}$  for relation symbols from  $\tau'$  and  $\tau''$  are as in  $\mathcal{P}'$  and  $\mathcal{P}''$  (and thus disjunction-free). The update formula for  $U$  is  $\phi_\delta^U \stackrel{\text{def}}{=} U$  and

$$\begin{aligned} \phi_\delta^{\tilde{Q}} &\stackrel{\text{def}}{=} (\phi_\delta^{Q'} \wedge \neg U) \vee (\phi_\delta^{Q''} \wedge U) \\ &\equiv (\phi_\delta^{Q'} \vee \phi_\delta^{Q''}) \wedge (\neg U \vee \phi_\delta^{Q''}) \wedge (\phi_\delta^{Q'} \vee U). \end{aligned}$$

The program  $\mathcal{P}'''$  is obtained from  $\tilde{\mathcal{P}}$  by removing disjunctions from  $\phi_\delta^{\tilde{Q}}$  using the method<sup>10</sup> from the proof of Lemma 3.7. For example, the first clause is replaced by  $\neg R_{\neg(Q' \vee Q'')}$  where  $R_{\neg(Q' \vee Q'')}$  is a fresh auxiliary relation symbol intended to be always interpreted by the result of the query  $\neg(\phi_\delta^{Q'} \vee \phi_\delta^{Q''})$ .

<sup>10</sup>This method cannot be used for DYNCQ and DYNFO<sup>^</sup>.

The update formula for  $R_{\neg(Q' \vee Q'')}$  after a modification  $\delta$  is  $\neg\phi_\delta^{Q'} \wedge \neg\phi_\delta^{Q''}$ ; it is disjunction-free since, by our assumption,  $\phi_\delta^{Q'}$  and  $\phi_\delta^{Q''}$  both consist of a single atom. This concludes the proof of (a).

The program  $\mathcal{P}'''$  for (b) and (c) is over schema  $\tau''' = \{Q'''\} \cup \tau' \cup \tau''$ . Again all update formulas of  $\mathcal{P}'''$  for relation symbols from  $\tau'$  and  $\tau''$  are as in  $\mathcal{P}'$  and  $\mathcal{P}''$  and

$$\phi_\delta^{Q'''} = \phi_\delta^{Q'} \wedge \phi_\delta^{Q''}.$$

The case distinction is delegated to the initialization mapping. Recall that the size of the domain is fixed when the auxiliary relations are initialized. The initialization mapping  $\text{INIT}'''$  is as follows. If  $|D| = 1$  then

$$\text{INIT}'''(R) = \begin{cases} \text{INIT}''(Q'') & \text{for } R = Q''', \\ D^k & \text{for } R \in \tau', \\ \text{INIT}''(R'') & \text{for } R \in \tau'' \end{cases}$$

If  $|D| \geq 2$  then

$$\text{INIT}'''(R) = \begin{cases} \text{INIT}'(Q') & \text{for } R = Q''', \\ \text{INIT}'(R') & \text{for } R \in \tau', \\ D^k & \text{for } R \in \tau'' \end{cases}$$

Thus  $\text{INIT}'''$  selects either  $\phi_\delta^{Q'}$  or  $\phi_\delta^{Q''}$ , depending on the size of the domain. If  $|D| = 1$  then  $\phi_\delta^{Q'}$  always evaluates to true whereas  $\phi_\delta^{Q''}$  yields the same value as in  $\mathcal{P}''$ , and vice versa for  $|D| \geq 2$ . As update formulas do not use negation, all relations in the program, that is initialized to “true” ( $\mathcal{P}'$  or  $\mathcal{P}''$ ) remain “full” throughout.<sup>11</sup> This concludes the proof of (b).  $\square$

It remains to prove that all queries over 0-ary relations can be maintained in DYNPROPCQ. 0-ary relations can either be true (containing the empty tuple) or false (not containing the empty tuple and thus being empty), thus 0-ary atoms are basically propositional variables. Queries on 0-ary databases are therefore basically families of Boolean functions, one for each domain size. Such queries are not very interesting from the perspective of databases, but we need to show the following lemma as we used it in the previous proof.

As quantification in queries on 0-ary databases is useless, every FO query can be expressed by a quantifier-free formula and therefore can be maintained in DYNPROP. The following lemma shows that this can be sharpened.

**Lemma 3.9.** *Every query on a 0-ary database can be maintained by a DYNPROPCQ-program.*

*Proof.* Let  $\tau_{\text{in}}$  be an input schema with 0-ary relation symbols  $A_1, \dots, A_k$ . Further let  $\mathcal{Q}_1, \dots, \mathcal{Q}_m$  be an enumeration of all  $m = 2^{2^k}$  many queries on  $\tau_{\text{in}}$ . We actually show that all of them can be maintained by one DYNPROPCQ-program  $\mathcal{P}$  with auxiliary schema  $\tau_{\text{aux}} = \{R_1, \dots, R_m\}$  maintaining  $\mathcal{Q}_i$  in  $R_i$ , for every

<sup>11</sup>This cannot be guaranteed for DYNUCQ<sup>+</sup>.

$i \in \{1, \dots, m\}$ . To this end, let  $\varphi_1, \dots, \varphi_m$  be propositional formulas over  $\tau_{\text{in}}$  such that  $\varphi_i$  expresses  $\mathcal{Q}_i$  and each  $\varphi_i$  is in conjunctive normal form. Without loss of generality, no clause contains  $A_l$  and  $\neg A_l$  for any  $A_l \in \tau_{\text{in}}$  and any  $\varphi_i$ . As  $\tau_{\text{aux}}$  contains a relation symbol, for every propositional formula over  $A_1, \dots, A_k$ , it contains, in particular, an auxiliary relation symbol  $R_C$ , for every disjunctive clause over  $A_1, \dots, A_k$ .

The update formulas for  $R_j$  after changing input relation  $A_l$  can be constructed as follows. Let  $\mathcal{C}$  be the set of clauses of  $\varphi_j$ , i.e.  $\varphi_j = \bigwedge_{C \in \mathcal{C}} C$ . We denote by  $\mathcal{C}_{A_l}^+$ ,  $\mathcal{C}_{A_l}^-$  and  $\mathcal{C}_{A_l}$  the subsets of  $\mathcal{C}$  whose clauses contain  $A_l$ ,  $\neg A_l$  and neither  $A_l$  nor  $\neg A_l$ , respectively.

If  $A_l$  becomes true by a modification then  $\varphi_j$  evaluates to true if all clauses in  $\mathcal{C}_{A_l}$  and all clauses  $C \setminus \{\neg A_l\}$  with  $C \in \mathcal{C}_{A_l}^-$  evaluated to true before the modification (clauses in  $\mathcal{C}_{A_l}^+$  will evaluate to true after enabling  $A_l$ ).

If  $A_l$  becomes false by a modification then  $\varphi_j$  evaluates to true if all clauses in  $\mathcal{C}_{A_l}$  and all clauses  $C \setminus \{A_l\}$  with  $C \in \mathcal{C}_{A_l}^+$  evaluated to true before the modification (clauses in  $\mathcal{C}_{A_l}^-$  will evaluate to true after disabling  $A_l$ ).

Therefore the update formulas for  $R_j$  after updating  $A_l$  can be defined as follows:

$$\begin{aligned}\phi_{\text{INS}_{A_l}}^{R_j} &\stackrel{\text{def}}{=} \bigwedge_{C \in \mathcal{C}_{A_l}} R_C \wedge \bigwedge_{C \in \mathcal{C}_{A_l}^-} R_{C \setminus \{\neg A_l\}} \\ \phi_{\text{DEL}_{A_l}}^{R_j} &\stackrel{\text{def}}{=} \bigwedge_{C \in \mathcal{C}_{A_l}} R_C \wedge \bigwedge_{C \in \mathcal{C}_{A_l}^+} R_{C \setminus \{A_l\}}\end{aligned}$$

The initialization is straightforward. The correctness of this construction can be proved by induction over the length of modification sequences.  $\square$

Finally we prove that  $\text{DYN}\exists^*\text{FO} = \text{DYN}\forall^*\text{FO}$ , and therefore that unions of conjunctive queries with negation coincide with  $\text{DYN}\forall^*\text{FO}$  in the dynamic setting. The proof uses the replacement technique to maintain the complements of the auxiliary relations used in the  $\text{DYN}\exists^*\text{FO}$ -program via  $\text{DYN}\forall^*\text{FO}$ -formulas. A small complication arises from the fact, that the query relation (and not its complement) has to be maintained. This is solved by ensuring that the update formulas of the query relation are atomic.

A slightly more general result can be shown.

**Lemma 3.10.** *Let  $\mathbb{Q}$  be an arbitrary quantifier prefix. A query can be maintained in  $\text{DYN}\mathbb{Q}\text{FO}$  if and only if it can be maintained in  $\text{DYN}\overline{\mathbb{Q}}\text{FO}$ .*

*Proof.* Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be an arbitrary dynamic  $\text{DYN}\mathbb{Q}\text{FO}$ -program over schema  $\tau$ . By Lemma 3.5 we can assume, without loss of generality, that the update formulas of  $Q$  are atomic. We construct a dynamic  $\text{DYN}\overline{\mathbb{Q}}\text{FO}$ -program  $\mathcal{P}'$  over schema  $\hat{\tau} \cup \{Q'\}$  where  $\hat{\tau}$  contains a  $k$ -ary relation symbol  $\hat{R}$  for every  $k$ -ary  $R \in \tau$ . The intention is that  $\hat{R}$  is always equal to the complement of  $R$ .

We denote by  $\phi[\tau \rightarrow \hat{\tau}]$  the formula obtained from  $\phi$  by replacing every atom  $S(\vec{z})$  in  $\phi$  by  $\neg \hat{S}(\vec{z})$ . Then the update formulas of  $\mathcal{P}'$  are obtained as

$\phi_{\delta}^{\widehat{R}} \stackrel{\text{def}}{=} \neg \phi_{\delta}^R[\tau \rightarrow \widehat{\tau}]$  for every  $\widehat{R} \in \widehat{\tau}$ . Observe that this formula can be easily transformed into an  $\overline{\mathbb{Q}}$ FO-formula. Further  $\phi_{\delta}^{Q'} = \neg \phi_{\delta}^{\widehat{Q}}$  which is a  $\overline{\mathbb{Q}}$ FO-formula since  $\phi_{\delta}^{\widehat{Q}}$  is quantifier-free. The initialization mapping of  $\mathcal{P}'$  is straightforward.  $\square$

Now Theorems 3.1, 3.2 and 3.4 follow immediately from Lemmata 3.6, 3.7, 3.8 and 3.10.

### 3.2 Separation results

The dynamic complexity classes DYNPROPCQ and DYNPROP have been separated already in [21, Lemma 7.4] using the non-empty-set query. Here we extend this result to the following theorem.

**Theorem 3.11.** (a) *The class DYNPROPCQ is a strict subclass of DYNPROP.*

(b) *The class DYNPROP is a strict subclass of DYNCQ.*

For the sake of completeness we repeat the proof idea for the separation of DYNPROPCQ and DYNPROP as used in [21].

*Proof sketch (of Theorem 3.11(a)).* Towards a contradiction assume that there is a DYNPROPCQ-program  $\mathcal{P} = (P, \text{INIT}, Q)$  that maintains the query  $Q$  defined by  $\exists x U(x)$ . By possibly introducing additional auxiliary relations we can guarantee that no variable occurs more than once in any atom of an update formula of  $\mathcal{P}$ .

The following observation is the key to the proof. Let us assume that there is a unary relation atom  $R(u)$  in the formula  $\phi_{\text{DEL}}^Q(u)$  for the 0-ary query relation  $Q$  and that  $\mathcal{S}$  is a state in which the relation  $U$  contains at least two elements  $a \neq b$ . Then, necessarily,  $R^{\mathcal{S}}$  contains both,  $a$  and  $b$ , as otherwise deletion of  $a$  or  $b$  would make  $\phi_{\text{DEL}}^Q$  false without  $U$  becoming empty. This observation can be generalized: if a relation  $R$  has *deletion depth*  $k$ , that is, the distance of  $R$  from  $Q$  in the deletion dependency graph is  $k$ , and  $U$  contains at least  $k + 1$  elements, then  $R$  must contain all *diverse* tuples over  $U$ , that is, tuples that consist of pairwise distinct elements from  $U$ .

Now the idea is as follows. Consider a state  $\mathcal{S}$  in which the set  $U$  contains  $m + 1$  elements, where  $m$  is the maximum (finite) deletion depth of any relation symbol in  $\mathcal{P}$ . By the observation above, all relations whose symbols are reachable from  $Q$  in the deletion dependency graph of  $\mathcal{P}$  contain all diverse tuples over  $U^{\mathcal{S}}$ . Thus, all relation atoms over diverse tuples from  $U^{\mathcal{S}}$  evaluate to true. It is easy to show by induction on the length of modification sequences that this property holds (with respect to  $U^{\mathcal{S}'}$ ) for all states  $\mathcal{S}'$  that can be obtained from  $\mathcal{S}$  by deleting elements from  $U^{\mathcal{S}}$ . In particular, it holds for any such state in which  $U^{\mathcal{S}'}$  contains only one element  $a$ . But then,  $\phi_{\text{DEL}}^Q(a)$  evaluates to true in  $\mathcal{S}'$  and thus  $Q$  remains true after deletion of  $a$ , the desired contradiction to the assumed correctness of  $\mathcal{P}$ .  $\square$

To prove the second statement of Theorem 3.11 we show that the class DYNQF (to be defined in a moment) is a subclass of DYNCQ. This proves Theorem 3.11(b) as it is known from [9] that DYNPROP is strictly contained in DYNQF.

We give an informal account of DYNQF before the actual definition and the statement of the theorem. The class DYNQF is an extension of DYNPROP, and was introduced by Hesse in [13]. Update formulas in DYNQF are quantifier-free (as in DYNPROP) but auxiliary functions may be used. Thus, while DYNPROP formulas can only access the inserted or deleted tuple  $\vec{a}$  and the currently updated tuple  $\vec{b}$  of an auxiliary relation, DYNQF update formulas can access further elements of the universe obtained by applying auxiliary functions to elements of  $\vec{a}$  and  $\vec{b}$ . Further, upon modification of the input database, auxiliary functions may be updated via update terms that may use functions and if-then-else-constructs.

The following formalization is adapted from [9] and [21]. We extend our definition of schemata to allow also function symbols. Until the end of this section, a *schema (or signature)*  $\tau$  consists of a set  $\tau_{\text{rel}}$  of relation symbols and a set  $\tau_{\text{fun}}$  of function symbols together with an arity function  $\text{Ar} : \tau_{\text{rel}} \cup \tau_{\text{fun}} \mapsto \mathbb{N}$ . A schema is *relational* if  $\tau_{\text{fun}} = \emptyset$ . A *database*  $\mathcal{D}$  of schema  $\tau$  with domain  $D$  is a mapping that assigns to every relation symbol  $R \in \tau_{\text{rel}}$  a relation of arity  $\text{Ar}(R)$  over  $D$  and to every  $k$ -ary function symbol  $f \in \tau_{\text{fun}}$  a  $k$ -ary function.

Next, we extend our definition of update programs to auxiliary schemas with functions. For updating auxiliary functions case distinctions are allowed in addition to the composition of function terms.

**Definition 4.** (Update term) *Update terms* are inductively defined as follows:

- (1) Every variable is an update term.
- (2) If  $f$  is a  $k$ -ary function symbol and  $t_1, \dots, t_k$  are update terms, then  $f(t_1, \dots, t_k)$  is an update term.
- (3) If  $\phi$  is a quantifier-free update formula (possibly using update terms) and  $t_1$  and  $t_2$  are update terms, then  $\text{ITE}(\phi, t_1, t_2)$  is an update term.

The semantics of update terms associates with every update term  $t$  and interpretation  $I = (\mathcal{S}, \beta)$ , where  $\mathcal{S}$  is a state and  $\beta$  a variable assignment, a value  $\llbracket t \rrbracket_I$  from  $S$ . The semantics of (1) and (2) is straightforward. If  $\mathcal{S} \models \phi$  holds, then  $\llbracket \text{ITE}(\phi, t_1, t_2) \rrbracket_I$  is  $\llbracket t_1 \rrbracket_I$ , otherwise  $\llbracket t_2 \rrbracket_I$ .

The extension of the notion of update programs for auxiliary schemas with function symbols is now straightforward. An update program still has an update formula  $\phi_\delta^R$  (possibly using update terms) for every relation symbol  $R \in \tau_{\text{aux}}$  and every abstract modification  $\delta$ . Furthermore, it has, for every abstract modification  $\delta$  and every function symbol  $f \in \tau_{\text{aux}}$ , an update term  $t_\delta^f(\vec{x}; \vec{y})$ . For a concrete modification  $\delta(\vec{a})$  it redefines  $f$  for each tuple  $\vec{b}$  by evaluating  $t_\delta^f(\vec{a}; \vec{b})$  in the current state.

**Definition 5.** (DYNQF) DYNQF is the class of queries maintainable by quantifier-free update programs with (possibly) auxiliary functions.

We remark that our definition of DYNQF is slightly stronger than the usual definition. Here we allow for using update terms in update formulas for relations whereas in [9] only terms are allowed. This strengthens the result we are aiming at.

Before continuing, we give a toy example designed to give an impression of the expressive power of DYNQF. For more examples we refer to [13] and [9].

**Example 4.** Consider the unary graph query  $Q(x)$  that returns all nodes  $a$  of a given graph  $G$  with maximal outdegree .

We construct a unary DYNQF-program  $\mathcal{P}$  that maintains  $Q$  in a unary relation denoted by the designated symbol  $Q$ . The program uses two unary functions  $\text{SUCC}$  and  $\text{PRED}$  that shall encode a successor and its corresponding predecessor relation on the domain. For simplicity, but without loss of generality, we therefore assume that the domain is of the form  $D = \{0, \dots, n-1\}$ . For every state  $\mathcal{S}$ , the function  $\text{SUCC}^{\mathcal{S}}$  is then the standard successor function on  $D$  (with  $\text{SUCC}^{\mathcal{S}}(n-1) = n-1$ ), and  $\text{PRED}^{\mathcal{S}}$  is the standard predecessor function (with  $\text{PRED}^{\mathcal{S}}(0) = 0$ ). Both functions are initialized accordingly. In the following we refer to *numbers* and mean the position of elements in  $\text{SUCC}$ . The program has a constants that represents the numbers 0 and 1.

The program  $\mathcal{P}$  maintains two unary functions  $\# \text{EDGES}$  and  $\# \text{NODES}$ . The function  $\# \text{EDGES}$  counts, for every node  $a$ , the number of outgoing edges of  $a$ ; more precisely  $\# \text{EDGES}(a) = b$  if and only if  $b$  is the number of outgoing edges of  $a$ . The function  $\# \text{NODES}$  counts, for every number  $a$ , the number of nodes with  $a$  outgoing edges; more precisely  $\# \text{NODES}(a) = b$  if and only if  $b$  is the number of nodes with  $a$  outgoing edges. A constant  $\text{MAX}$  shall always point to the number  $i$  such that  $i$  is the maximal number of outgoing edges from some node in the current graph.

When inserting an outgoing edge  $(u, v)$  for a node  $u$  that already has  $a$  outgoing edges, the counter  $\# \text{EDGES}$  of  $u$  is incremented from  $a$  to  $a+1$  and all other edge-counters remain unchanged. The counter  $\# \text{NODES}$  of  $a$  is decremented, the counter of  $a+1$  is incremented, and all other node-counters remain unchanged. The number  $\text{MAX}$  increases if, before the insertion,  $u$  was a node with maximal number of outgoing edges. This yields the following update terms:

$$\begin{aligned}
t_{\text{INS } E}^{\# \text{EDGES}}(u, v; x) &\stackrel{\text{def}}{=} \text{ITE}\left(\neg E(u, v) \wedge x = u, \text{SUCC}(\# \text{EDGES}(x)), \# \text{EDGES}(x)\right) \\
t_{\text{INS } E}^{\# \text{NODES}}(u, v; x) &\stackrel{\text{def}}{=} \text{ITE}\left(\neg E(u, v) \wedge x = \# \text{EDGES}(u), \text{PRED}(\# \text{NODES}(x)), \right. \\
&\quad \left. \text{ITE}(\neg E(u, v) \wedge x = \text{SUCC}(\# \text{EDGES}(u)), \text{SUCC}(\# \text{NODES}(x)), \right. \\
&\quad \left. \# \text{NODES}(x))\right) \\
t_{\text{INS } E}^{\text{MAX}}(u, v) &\stackrel{\text{def}}{=} \text{ITE}\left(\text{MAX} = \# \text{EDGES}(u) \wedge \neg E(u, v), \text{SUCC}(u), \text{MAX}\right)
\end{aligned}$$

The update formula for the designated query symbol  $Q$  is as follows:

$$\phi_{\text{INS } E}^Q(u, v; x) \stackrel{\text{def}}{=} t_{\text{INS } E}^{\# \text{EDGES}}(u, v; x) = t_{\text{INS } E}^{\text{MAX}}(u, v)$$

The update terms for deletions are very similar:

$$\begin{aligned} t_{\text{DEL } E}^{\# \text{EDGES}}(u, v; x) &\stackrel{\text{def}}{=} \text{ITE}\left(E(u, v) \wedge x = u, \text{PRED}(\# \text{EDGES}(x)), \# \text{EDGES}(x)\right) \\ t_{\text{DEL } E}^{\# \text{NODES}}(u, v; x) &\stackrel{\text{def}}{=} \text{ITE}\left(E(u, v) \wedge x = \# \text{EDGES}(u), \text{PRED}(\# \text{NODES}(x)), \right. \\ &\quad \left. \text{ITE}(E(u, v) \wedge x = \text{PRED}(\# \text{EDGES}(u)), \text{SUCC}(\# \text{NODES}(x)), \right. \\ &\quad \left. \# \text{NODES}(x))\right) \\ t_{\text{DEL } E}^{\text{MAX}}(u, v) &\stackrel{\text{def}}{=} \text{ITE}\left(\text{MAX} = \# \text{EDGES}(u) \wedge E(u, v) \wedge \# \text{NODES}(\text{MAX}) = 1, \right. \\ &\quad \left. \text{PRED}(\text{MAX}), \text{MAX}\right) \end{aligned}$$

The update formula for the designated query symbol  $Q$  under deletion is as follows:

$$\phi_{\text{DEL } E}^Q(u, v; x) \stackrel{\text{def}}{=} t_{\text{DEL } E}^{\# \text{EDGES}}(u, v; x) = t_{\text{DEL } E}^{\text{MAX}}(u, v)$$

□

We aim at the following theorem.

**Theorem 3.12.** *DYNQF is contained in DYNCQ.*

Thanks to Theorem 3.2 it suffices to show that DYNQF is contained in DYNUCQ. The idea of the proof of Theorem 3.12 is to simulate auxiliary functions by auxiliary relations with the help of existential quantifiers in a relatively straightforward way. However, some care is necessary in order to remove ITE-conditions and negations. We highlight the idea of the proof by the following example.

**Example 5.** Consider a DYNQF-program  $\mathcal{P}$  that contains the following update term  $t_{\text{INSE}}^f$  for a unary function  $f$  and update formula  $\phi_{\text{INSE}}^R$  for a unary relation  $R$ :

$$\begin{aligned} t_{\text{INSE}}^f(u, v; x) &\stackrel{\text{def}}{=} f(\text{ITE}(R(x), f(x), u)) \\ \phi_{\text{INSE}}^R(u, v; x) &\stackrel{\text{def}}{=} \neg R(x) \wedge S(f(x), \text{ITE}(\neg R(\text{ITE}(S(u, v), u, x)), f(x), u)) \end{aligned}$$

As a first step towards the construction of an equivalent DYNCQ-program, we remove negations by maintaining for every relation  $T$  its complement in an auxiliary relation  $\widehat{T}$ , for example:

$$\begin{aligned} \phi_{\text{INSE}}^R(u, v; x) &\stackrel{\text{def}}{=} \widehat{R}(x) \wedge S(f(x), \text{ITE}(\widehat{R}(\text{ITE}(S(u, v), u, x)), f(x), u)) \\ \phi_{\text{INSE}}^{\widehat{R}}(u, v; x) &\stackrel{\text{def}}{=} R(x) \vee \widehat{S}(f(x), \text{ITE}(\widehat{R}(\text{ITE}(S(u, v), u, x)), f(x), u)) \end{aligned}$$



The crucial step in the construction of an equivalent DYNCQ-program is to simulate the function  $f$  by a binary relation  $R_f$  with the intention that  $R_f$  stores all tuples of the form  $(a, f(a))$ . Then appearances of  $f$  as well as of ITE can be removed. The complement relations obtained in the preprocessing step above are also needed in order to remove ITE-conditions.

The update formula  $\phi_{\text{INSE}}^{R_f}(u, v; x, y)$  for  $R_f$  is obtained by 'outsourcing' the computation of the ITE-value:

$$\phi_{\text{INSE}}^{R_f}(u, v; x, y) \stackrel{\text{def}}{=} \exists z \left( R_f(z, y) \wedge \underbrace{\left( (R(x) \wedge R_f(x, z)) \vee (\widehat{R}(x) \wedge z = u) \right)}_{z = \text{ITE}(R(x), f(x), u)} \right)$$

An update formula for  $R$  that uses  $R_f$  instead of  $f$  is obtained similarly:

$$\begin{aligned} \phi_{\text{INSE}}^R(u, v; x) &\stackrel{\text{def}}{=} \widehat{R}(x) \wedge \exists z_1 \exists z_2 \exists z_3 \left( S(z_1, z_2) \wedge R_f(x, z_1) \right. \\ &\quad \left. \wedge \left( (\widehat{R}(z_3) \wedge R_f(x, z_2)) \vee (R(z_3) \wedge z_2 = u) \right) \right\} \text{ITE}(\widehat{R}(\dots), \cdot, \cdot) \\ &\quad \wedge \left( (S(u, v) \wedge z_3 = u) \vee (\widehat{S}(u, v) \wedge z_3 = x) \right) \left\} \text{ITE}(S(\dots), \cdot, \cdot) \right) \end{aligned}$$

Observe that only relation symbols from the original DYNQF-program are needed in negated form.

*Proof (of Theorem 3.12).* Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a DYNQF-program over schema  $\tau = \tau_{\text{rel}} \cup \tau_{\text{fun}}$ . We assume, without loss of generality, that  $\mathcal{P}$  is in negation normal form. Further we assume, as in Lemma 3.6, that the input relations have update formulas as well.

We prove that there is a DYNUCQ-program  $\mathcal{P}''$  equivalent to  $\mathcal{P}$ . Then, by Theorem 3.2, there is an equivalent DYNCQ-program.

As a preparation step we construct, from  $\mathcal{P}$ , a DYNQF-program  $\mathcal{P}'$  over schema  $\tau' \stackrel{\text{def}}{=} \tau_{\text{rel}} \cup \widehat{\tau}_{\text{rel}} \cup \{\widehat{=}\} \cup \tau_{\text{fun}}$  where  $\widehat{\tau}_{\text{rel}}$  has, for every  $R \in \tau$ , a relation symbol  $\widehat{R}$  intended to contain the complement of  $R$  and  $\widehat{=}$  contains the complement of the relation  $=$ . This can be achieved as in the proof of Lemma 3.6.

From  $\mathcal{P}'$  we construct a DYNUCQ-program  $\mathcal{P}''$  over schema  $\tau'' \stackrel{\text{def}}{=} \tau_{\text{rel}} \cup \widehat{\tau}_{\text{rel}} \cup \tau_F$  where  $\tau_F$  contains a  $(k+1)$ -ary relation symbol  $R_f$  for every  $k$ -ary function symbol  $f \in \tau_{\text{fun}}$ . The intention is that  $R_f$  simulates  $f$  in the sense that  $(\vec{a}, b) \in R_f^{\mathcal{S}'}$  if and only if  $f^{\mathcal{S}''}(\vec{a}) = b$  in states  $\mathcal{S}'$  and  $\mathcal{S}''$  reached in  $\mathcal{P}'$  and  $\mathcal{P}''$  by the same modification sequence. The initialization of  $R_f$  can be obtained easily from the initialization of  $f$ .

We say that two states  $\mathcal{S}'$  and  $\mathcal{S}''$  over  $\tau'$  and  $\tau''$  *correspond*, if (1) the condition  $(\vec{a}, b) \in R_f^{\mathcal{S}'}$  if and only if  $f^{\mathcal{S}''}(\vec{a}) = b$  is satisfied, and (2)  $R^{\mathcal{S}'} = R^{\mathcal{S}''}$  for all  $R \in \tau_{\text{rel}} \cup \widehat{\tau}_{\text{rel}}$ .

We explain next how to update relations from  $\tau_F$ . To this end, we will define CQ-formulas  $\varphi_t(\vec{x}, z)$  and  $\varphi_\phi(\vec{x})$  over  $\tau''$ , for every update term  $t(\vec{x})$  and every update formula  $\phi(\vec{x})$  over  $\tau'$ , such that the following conditions are satisfied for all corresponding states  $\mathcal{S}'$ ,  $\mathcal{S}''$ , all tuples  $\vec{a}$  and all elements  $b$ :

- $\mathcal{S}'', (\vec{a}, b) \models \varphi_t(\vec{x}, z)$  if and only if  $t^{\mathcal{S}'}(\vec{a}) = b$ , and

- $\mathcal{S}'', \vec{a} \models \varphi_\phi(\vec{x})$  if and only if  $\mathcal{S}', \vec{a} \models \phi(\vec{x})$

Then the update formulas in  $\mathcal{P}''$  after a modification  $\delta$  can be defined as follows. For every  $R_f \in \tau_F$ , define the update formula as  $\phi_\delta^{R_f}(\vec{u}; \vec{x}, y) \stackrel{\text{def}}{=} \varphi_t(\vec{u}, \vec{x}, y)$  where  $t$  is the update term for  $f \in \tau_{\text{fun}}$  in  $\mathcal{P}'$ . For every  $R \in \tau_{\text{rel}} \cup \widehat{\tau}_{\text{rel}}$  define the update formula as  $\phi_\delta^R \stackrel{\text{def}}{=} \varphi_\phi$  where  $\phi$  is the update formula of  $R$  in  $\mathcal{P}'$ . An easy induction shows that  $\mathcal{P}'$  and  $\mathcal{P}''$  yield corresponding states when the same modification sequence is applied. This proves the claim.

It remains to define the CQ-formulas  $\varphi_t(\vec{x}, z)$  and  $\varphi_\phi(\vec{x})$  for every update term  $t(\vec{x})$  and every formula  $\phi(\vec{x})$ . Those formulas are defined inductively as follows:

- (a) If  $t(\vec{x}) = y$  for some variable  $y$  occurring in  $\vec{x}$ , then

$$\varphi_t(\vec{x}, z) \stackrel{\text{def}}{=} y = z.$$

- (b) If  $t(\vec{x}) = f(t_1(\vec{x}_1), \dots, t_k(\vec{x}_k))$  with  $\vec{x}_i \subseteq \vec{x}$ , then

$$\varphi_t(\vec{x}, z) \stackrel{\text{def}}{=} \exists z_1 \dots \exists z_k \left( R_f(z_1, \dots, z_k, z) \wedge \bigwedge_i \varphi_{t_i}(\vec{x}_i, z_i) \right).$$

- (c) If  $t(\vec{x}) = \text{ITE}(\phi(\vec{y}), t_1(\vec{x}_1), t_2(\vec{x}_2))$  with  $\vec{y}, \vec{x}_1, \vec{x}_2 \subseteq \vec{x}$ , quantifier-free update formula  $\phi$  and update terms  $t_1, t_2$ , then

$$\varphi_t(\vec{x}, z) \stackrel{\text{def}}{=} (\varphi_\phi(\vec{y}) \wedge \varphi_{t_1}(\vec{x}_1, z)) \vee (\varphi_{\neg\phi}(\vec{y}) \wedge \varphi_{t_2}(\vec{x}_2, z)).$$

- (d) If  $\phi(\vec{x})$  contains the maximal<sup>12</sup> update terms  $t_1(\vec{x}_1), \dots, t_k(\vec{x}_k)$  then let

$$\varphi_\phi(\vec{x}) \stackrel{\text{def}}{=} \exists z_1 \dots \exists z_k \left( \phi' \wedge \bigwedge_i \varphi_{t_i}(\vec{x}_i, z_i) \right)$$

where  $\phi'$  is obtained from  $\phi$  by replacing  $t_i$  by  $z_i$ , transforming the resulting formula into negation normal form and then replacing every literal of the form  $\neg R(s_1, \dots, s_l)$  by  $\widehat{R}(s_1, \dots, s_l)$ .

Observe that the formula  $\phi$  in (d) contains only relation symbols from  $\tau_{\text{rel}} \cup \widehat{\tau}_{\text{rel}}$ , and therefore no relation symbols from  $\tau_{\text{fun}}$  need to be replaced in  $\phi'$ . The correctness of this construction can be proved inductively.  $\square$

## 4 $\Delta$ -semantics

So far we considered a semantics where the new version of the auxiliary relations is redefined, after each modification, from scratch by formulas that are evaluated

<sup>12</sup>Here, a term  $t_i$  is maximal if it is not contained in another update term.

on the structure with the current auxiliary relations. We refer to this as *absolute semantics* in the following.

However, in the context of view maintenance, one usually expects only few auxiliary tuples to change after a modification. Therefore it is common to express the new version of the auxiliary relations in terms of the current relations and some “Delta”, that is, a (small) relation  $R^+$  of tuples to be inserted into  $R$  and a (small) relation  $R^-$  of tuples to be removed from  $R$  (with  $R^+ \cap R^- = \emptyset$ ). The updated auxiliary relation  $R'$  is then defined by

$$R' \stackrel{\text{def}}{=} (R \cup R^+) \setminus R^-.$$

We refer to this semantics as  $\Delta$ -*semantics*. This is the semantics usually considered in view maintenance. As already stated in the introduction, absolute and  $\Delta$ -semantics can only be different if the underlying update language is not closed under Boolean operations.

Next we formalize  $\Delta$ -semantics via  $\Delta$ -update programs which provide formulas defining the relations  $R^+$  and  $R^-$ , for every auxiliary relation  $R$ .

**Definition 6.** ( $\Delta$ -Update program) A  $\Delta$ -update program  $\mathcal{P}$  over dynamic schema  $(\tau_{\text{in}}, \tau_{\text{aux}})$  is a set of first-order formulas (called  $\Delta$ -update formulas in the following) that contains, for every  $R \in \tau_{\text{aux}}$  and every  $\delta \in \{\text{INS}_S, \text{DEL}_S\}$  with  $S \in \tau_{\text{in}}$ , two formulas  $\phi_\delta^{R^+}(\vec{u}; \vec{x})$  and  $\phi_\delta^{R^-}(\vec{u}; \vec{x})$  over the schema  $\tau$  where  $\vec{u}$  and  $S$  have the same arity,  $\vec{x}$  and  $R$  have the same arity, and  $\phi_\delta^{R^+} \wedge \phi_\delta^{R^-}$  is unsatisfiable.

The *semantics of  $\Delta$ -update programs* is as follows. For a modification  $\delta = \delta(\vec{a})$  and program state  $\mathcal{S} = (D, \mathcal{I}, \mathcal{A})$  we denote by  $P_\delta(\mathcal{S})$  the state  $(D, \delta(\mathcal{I}), \mathcal{A}')$ , where the relations  $R'$  of  $\mathcal{A}'$  are defined by

$$R' \stackrel{\text{def}}{=} \left( R \cup \{ \vec{b} \mid \mathcal{S} \models \phi_\delta^{R^+}(\vec{a}; \vec{b}) \} \right) \setminus \{ \vec{b} \mid \mathcal{S} \models \phi_\delta^{R^-}(\vec{a}; \vec{b}) \}.$$

The effect of a modification sequence on a state, dynamic  $\Delta$ -programs and so on are defined like their counterparts in absolute semantics except that  $\Delta$ -update programs are used instead of update programs.

**Definition 7.** ( $\Delta$ -DYN $\mathcal{C}$ ) For a class  $\mathcal{C}$  of formulas, let  $\Delta\text{-DYN}\mathcal{C}$  be the class of all dynamic queries that can be maintained by dynamic  $\Delta$ -programs with formulas from  $\mathcal{C}$  and arbitrary initialization mapping.

We note that the definitions above do *not* require that  $R^+ \cap R = \emptyset$  or  $R^- \subseteq R$ , that is,  $R^+$  might contain tuples that are already in  $R$ , and  $R^-$  might contain tuples that are not in  $R$ . However, in all proofs below, we construct only  $\Delta$ -update formulas that guarantee these additional properties. As a consequence, for the considered fragments, the expressive power is independent of this difference.

The goal of this section is to prove the remaining results of Figure 1, that is, the collapse results depicted in the right part of the figure and the correspondences between absolute semantics and  $\Delta$ -semantics.

**Theorem 4.1.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNUCQ}^\neg$ .
- (b)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNUCQ}$ .
- (c)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNQCQ}^\neg$ .
- (d)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNQCQ}$ .
- (e)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYN}\exists^*\text{FO}$ .
- (f)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYN}\forall^*\text{FO}$ .

**Theorem 4.2.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in  $\text{DYNUCQ}^\neg$ .
- (b)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNUCQ}^\neg$ .

The technique used for removing unions from dynamic unions of conjunctive queries under  $\Delta$ -semantics can be used to obtain a  $\Delta\text{-DYNFO}^\wedge$  normal form for  $\Delta\text{-DYNFO}$ -programs.

**Theorem 4.3.** *Let  $\mathcal{Q}$  be a query. Then the following statements are equivalent:*

- (a)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNFO}$ .
- (b)  $\mathcal{Q}$  can be maintained in  $\Delta\text{-DYNFO}^\wedge$ .

We state some basic facts about dynamic programs with  $\Delta$ -semantics before proving those theorems. The following lemma establishes the obvious fact that the absolute semantics and  $\Delta$ -semantics coincide in expressive power for dynamic classes closed under boolean operations. We observe that the proof does not work for (extensions of) conjunctive queries. Later we will see how to extend the result to conjunctive queries.

**Lemma 4.4.** *Let  $\mathcal{C}$  be some fragment of first-order logic closed under the boolean operations  $\{\vee, \wedge, \neg\}$ . Then for every query  $\mathcal{Q}$  the following are equivalent:*

- (a) *There is a  $\text{DYN}\mathcal{C}$ -program that maintains  $\mathcal{Q}$ .*
- (b) *There is a  $\Delta\text{-DYN}\mathcal{C}$ -program that maintains  $\mathcal{Q}$ .*

*Proof.* From an  $\text{DYN}\mathcal{C}$ -update formula  $\phi_\delta^R$ , the  $\Delta\text{-DYN}\mathcal{C}$ -update formulas are defined as follows:

$$\begin{aligned}\phi_\delta^{R^+}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \phi_\delta^R(\vec{u}; \vec{x}) \wedge \neg R(\vec{x}) \\ \phi_\delta^{R^-}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \neg \phi_\delta^R(\vec{u}; \vec{x}) \wedge R(\vec{x})\end{aligned}$$

From a  $\Delta\text{-DYN}\mathcal{C}$ -update formulas  $\phi_\delta^{R^+}$  and  $\phi_\delta^{R^-}$ , an  $\text{DYN}\mathcal{C}$ -update formula is obtained via

$$\phi_\delta^R(\vec{u}; \vec{x}) \stackrel{\text{def}}{=} (R(\vec{x}) \vee \phi_\delta^{R^+}(\vec{u}; \vec{x})) \wedge \neg \phi_\delta^{R^-}(\vec{u}; \vec{x}).$$

□

Removing negations in dynamic programs with  $\Delta$ -semantics is straightforward using the replacement technique, since the complement  $\widehat{R}$  of an auxiliary relation  $R$  can be maintained by exchanging the formulas  $\phi_\delta^{R^+}$  and  $\phi_\delta^{R^-}$ . Observe that in contrast to absolute semantics this works for arbitrary query classes, even if they are not closed under complementation, and in particular for (extensions of) conjunctive queries.

**Lemma 4.5.** *Let  $\mathcal{C}$  be some fragment of first-order logic. If a query  $\mathcal{Q}$  can be maintained in  $\Delta$ -DYN $\mathcal{C}$  then  $\mathcal{Q}$  can be maintained in negation-free  $\Delta$ -DYN $\mathcal{C}$ .*

*Proof.* The idea is again to maintain the complements for auxiliary relations. Given a dynamic  $\Delta$ -program  $\mathcal{P}$  over schema  $\tau$  we construct a dynamic  $\Delta$ -program  $\mathcal{P}'$  over schema  $\tau \cup \widehat{\tau}$  where  $\widehat{\tau}$  contains, for every  $k$ -ary relation symbol  $R \in \tau$ , a fresh  $k$ -ary relation symbol  $\widehat{R}$  with the intention that  $\widehat{R}$  always stores the complement of  $R$ .

The update formulas for  $R \in \tau$  are as in  $\mathcal{P}$ . For a relation symbol  $R \in \tau$  let  $\phi_\delta^{R^+}(\vec{u}; \vec{x})$  and  $\phi_\delta^{R^-}(\vec{u}; \vec{x})$  be the update formulas of  $R$ . Then the update formulas for  $\widehat{R}$  can be defined as follows:

$$\begin{aligned}\phi_\delta^{\widehat{R}^+}(\vec{u}; \vec{x}) &= \phi_\delta^{R^-}(\vec{u}; \vec{x}) \\ \phi_\delta^{\widehat{R}^-}(\vec{u}; \vec{x}) &= \phi_\delta^{R^+}(\vec{u}; \vec{x})\end{aligned}$$

From  $\mathcal{P}'$ , a negation-free dynamic  $\Delta$ -program  $\mathcal{P}''$  can be constructed by replacing, for all  $R \in \tau$ , all occurrences of  $\neg R(\vec{x})$  in update formulas of  $\mathcal{P}'$  by  $\widehat{R}(\vec{x})$ . We omit the obvious proof of correctness.  $\square$

We now turn towards proving the main results of this section. We first prove Theorem 4.2. Afterwards we use the connection between absolute and  $\Delta$ -semantics that it establishes as well as the adaption of Lemma 3.8 to  $\Delta$ -semantics to prove the characterization of conjunctive queries with  $\Delta$ -semantics.

The only-if-direction of Theorem 4.2 can be generalized to arbitrary quantifier prefixes. It is open whether the if-direction generalizes as well.

**Lemma 4.6.** *Let  $\mathbb{Q}$  be an arbitrary quantifier prefix. If a query can be maintained in DYN $\mathbb{Q}$ FO then it can be maintained in  $\Delta$ -DYN $\mathbb{Q}$ FO as well.*

*Proof.* Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a DYN $\mathbb{Q}$ FO-program with schema  $\tau$ . By Lemma 3.5 we can assume, without loss of generality, that the update formulas of  $Q$  are atomic. We construct a dynamic  $\Delta$ -DYN $\mathbb{Q}$ FO-program  $\mathcal{P}' = (P', \text{INIT}', Q')$ .

The main challenge is to design update formulas of the kind  $\phi_\delta^{R^-}$  without being able to complement the given update formulas because this would lead to  $\overline{\mathbb{Q}}$ FO-formulas (additionally, the disjointness requirement for formulas  $\phi_\delta^{R^+}$  needs to be ensured).

The basic idea is to use two copies of the auxiliary relations, both alternating between empty and useful states, such that one copy is useful for even steps and the other one for odd steps. More precisely, for every auxiliary relation  $R$  used by  $\mathcal{P}$ , the program  $\mathcal{P}'$  uses two auxiliary relations  $R_{\text{even}}$  and  $R_{\text{odd}}$  with the

intention that after an even sequence of modifications  $R_{\text{even}}$  stores the content of  $R$  after the same sequence of modifications while  $R_{\text{odd}}$  is empty. After an odd sequence of modifications  $R_{\text{even}}$  is empty while  $R_{\text{odd}}$  stores the content of  $R$ .

Then, for an even modification, the relation  $R_{\text{even}}^+$  can be simply expressed as in absolute semantics (using “odd” relations) and  $R_{\text{even}}^-$  is empty. For an odd modification  $R_{\text{even}}^-$  can be simply chosen as  $R_{\text{even}}$  and  $R_{\text{even}}^+$  is empty. Similarly for  $R_{\text{odd}}$ .

In the following we give a precise construction of  $\mathcal{P}'$  over schema  $\tau_{\text{even}} \cup \tau_{\text{odd}} \cup \{\text{ODD}, Q'\}$  where ODD is a boolean relation symbol, and  $\tau_{\text{even}}$  and  $\tau_{\text{odd}}$  contain, for every  $k$ -ary relation symbol  $R \in \tau$ , a  $k$ -ary relation symbol  $R_{\text{even}}$  and  $R_{\text{odd}}$ , respectively. The relation ODD is used to store the parity of the number of modifications performed so far.

Let  $\phi_\delta^R$  be the update formula of  $R \in \tau$  for a modification  $\delta$  in the dynamic program  $\mathcal{P}$ . Denote by  $\phi_\delta^R[\tau \rightarrow \tau_{\text{even}}]$  the formula obtained from  $\phi_\delta^R$  by replacing every atom  $S(\vec{x})$  with  $S \in \tau$  by  $S_{\text{even}}(\vec{x})$ . Analogously for  $\phi_\delta^R[\tau \rightarrow \tau_{\text{odd}}]$ . Now, the update formulas for  $R_{\text{odd}}$  and  $R_{\text{even}}$  are as follows:

$$\begin{aligned}\phi_\delta^{R_{\text{odd}}}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \neg\text{ODD} \wedge \phi_\delta^R[\tau \rightarrow \tau_{\text{even}}](\vec{u}; \vec{x}) \\ \phi_\delta^{R_{\text{odd}}}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \text{ODD} \wedge R_{\text{odd}}(\vec{x}) \\ \phi_\delta^{R_{\text{even}}}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \text{ODD} \wedge \phi_\delta^R[\tau \rightarrow \tau_{\text{odd}}](\vec{u}; \vec{x}) \\ \phi_\delta^{R_{\text{even}}}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \neg\text{ODD} \wedge R_{\text{even}}(\vec{x})\end{aligned}$$

Observe that all those formulas can be easily converted into QFO-formulas. The boolean auxiliary relation ODD can be updated straightforwardly.

Now, since the update formulas of  $Q$  in  $\mathcal{P}$  are quantifier-free, the relation  $Q'$  can be updated with the following quantifier-free update formulas:

$$\begin{aligned}\phi_\delta^{Q'^+}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \phi_\delta^Q(\vec{u}; \vec{x}) \wedge \\ &\quad \neg \left( (\text{ODD} \wedge Q_{\text{odd}}(\vec{x})) \vee (\neg\text{ODD} \wedge Q_{\text{even}}(\vec{x})) \right) \\ \phi_\delta^{Q'^-}(\vec{u}; \vec{x}) &\stackrel{\text{def}}{=} \neg\phi_\delta^Q(\vec{u}; \vec{x}) \wedge \\ &\quad \left( (\text{ODD} \wedge Q_{\text{odd}}(\vec{x})) \vee (\neg\text{ODD} \wedge Q_{\text{even}}(\vec{x})) \right)\end{aligned}$$

The initialization mapping of  $\mathcal{P}'$  is straightforward. Every  $R_{\text{even}} \in \tau_{\text{even}}$  is initialized with  $\text{INIT}(R)$ . All  $R_{\text{odd}} \in \tau_{\text{odd}}$  are initialized with the empty relation. The relation ODD is initialized with  $\perp$ , and  $Q'$  is initialized with  $\text{INIT}(Q)$ .  $\square$

**Lemma 4.7.** (a) *If a query can be maintained in  $\Delta\text{-DYNUCQ}^-$  then it can be maintained in  $\text{DYNUCQ}^-$  as well.*

(b) *If a query can be maintained in  $\Delta\text{-DYN}\forall^*\text{FO}$  then it can be maintained in  $\text{DYN}\forall^*\text{FO}$  as well.*

We note that the first statement could equally be expressed in terms of  $\Delta\text{-DYN}\exists^*\text{FO}$  and  $\text{DYN}\exists^*\text{FO}$ .

*Proof.* We only prove (a), the proof of (b) is analogous. Let  $\mathcal{P} = (P, \text{INIT}, Q)$  be a dynamic  $\Delta\text{-DYNUCQ}^-$ -program over schema  $\tau$ . By Lemma 4.5 we can assume, without loss of generality, that the update formulas of  $\mathcal{P}$  are negation-free. For ease of presentation we assume that the input schema contains a single binary relation symbol  $E$ .

We construct an equivalent  $\text{DYNUCQ}^-$ -program  $\mathcal{P}'$  using the following idea. Consider some update formulas  $\phi_\delta^{R^+}(\vec{u}; \vec{x})$  and  $\phi_\delta^{R^-}(\vec{u}; \vec{x})$  of a relation  $R \in \tau$  for a modification  $\delta$  in  $\mathcal{P}$ . The naïve translation into a  $\text{DYNFO}$ -update formula  $\phi_\delta^R(\vec{u}; \vec{x})$  yields the formula

$$\phi_\delta^R(\vec{u}; \vec{x}) = (R(\vec{x}) \vee \phi_\delta^{R^+}(\vec{u}; \vec{x})) \wedge \neg \phi_\delta^{R^-}(\vec{u}; \vec{x})$$

which is possibly non- $\text{UCQ}^-$  due to  $\neg \phi_\delta^{R^-}(\vec{u}; \vec{x})$ . Therefore,  $\mathcal{P}'$  maintains a relation  $R_\delta^-$  that contains all tuples  $(\vec{a}, \vec{b})$  such that  $\vec{a}$  would be removed from  $R$  after applying the modification  $\delta(\vec{b})$ . Those relations are maintained using the squirrel technique. The dynamic program  $\mathcal{P}'$  is over schema  $\tau \cup \tau_\Delta$  where  $\tau_\Delta$  contains a  $(k+2)$ -ary relation symbol  $R_\delta^- \in \tau$  for every  $k$ -ary relation symbol  $R \in \tau$  and every modification  $\delta \in \{\text{INS}, \text{DEL}\}$  of the input relation  $E$ .

The update formula for a relation symbol  $R \in \tau$  is

$$\phi_\delta^R(\vec{u}; \vec{x}) \stackrel{\text{def}}{=} (R(\vec{x}) \vee \phi_\delta^{R^+}(\vec{u}; \vec{x})) \wedge \neg R_\delta^-(\vec{u}, \vec{x}).$$

This formula can be translated into an existential formula in a straightforward manner.

For updating a relation  $R_{\delta_1}^-$  after a modification  $\delta_0$ , the update formula  $\phi_{\delta_1}^{R^-}$  for  $R^-$  is used. However, since  $R_{\delta_1}^-$  shall store tuples that have to be deleted after applying  $\delta_1$ , the formula  $\phi_{\delta_1}^{R^-}$  has to be adapted to use the content of relation symbols  $S \in \tau$  after modification  $\delta_0$  (instead, as usual, the content from before the modification). For this purpose relation symbols  $S \in \tau$  in  $\phi_{\delta_1}^{R^-}$  need to be replaced by their update formulas as defined above.

The update formula for  $R_{\delta_1}^-$  is

$$\phi_{\delta_0}^{R_{\delta_1}^-}(\vec{u}_0; \vec{u}_1, \vec{x}) \stackrel{\text{def}}{=} \phi_{\delta_0}^{R_{\delta_1}^-}[\tau \rightarrow \phi^\tau](\vec{u}_0; \vec{u}_1, \vec{x})$$

where  $\phi_{\delta_0}^{R_{\delta_1}^-}[\tau \rightarrow \phi^\tau](\vec{u}_0; \vec{u}_1, \vec{x})$  is obtained from  $\phi_{\delta_1}^{R^-}(\vec{u}; \vec{x})$  by replacing every atom  $S(\vec{z})$  by  $\phi_{\delta_0}^S(\vec{u}_0; \vec{z})$ , as constructed above. Since by our initial assumption,  $\phi_{\delta_1}^{R^-}$  itself is an existential formula without negation and all update formulas  $\phi_{\delta_0}^S$  for  $S \in \tau$  are existential, the formula  $\phi_{\delta_0}^{R_{\delta_1}^-}$  can be easily converted into an existential formula as well.  $\square$

The following example illustrates the construction of Lemma 4.7.

**Example 6.** Consider the following negation-free  $\Delta$ -update formulas for a relation symbol  $R$ :

$$\begin{aligned}\phi_{\text{INS}}^{R^+}(u; x) &= \exists y (R(y) \wedge S(u, x)) \\ \phi_{\text{INS}}^{R^-}(u; x) &= \exists y (U(x) \vee (R(y) \wedge S(y, u))) \\ \phi_{\text{DEL}}^{R^+}(u; x) &= \exists y U(y) \\ \phi_{\text{DEL}}^{R^-}(u; x) &= \exists y \exists z (S(x, z) \wedge S(y, u))\end{aligned}$$

Then the construction from the previous Lemma 4.7 yields the following update formulas for  $R$  and  $R_{\delta_1}^-$  which can be easily translated into UCQ $^-$ -formulas:

$$\begin{aligned}\phi_{\text{INS}}^R(u; x) &= (R(x) \vee \phi_{\text{INS}}^{R^+}(u; x)) \wedge \neg R_{\text{INS}}^-(u, x) \\ \phi_{\text{DEL}}^R(u; x) &= (R(x) \vee \phi_{\text{DEL}}^{R^+}(u; x)) \wedge \neg R_{\text{DEL}}^-(u, x) \\ \phi_{\text{INS}}^{R^-}(u_0; u_1, x) &= \exists y (\phi_{\text{INS}}^U(u_0; x) \vee (\phi_{\text{INS}}^R(u_0; y) \wedge \phi_{\text{INS}}^S(u_0; y, u_1))) \\ \phi_{\text{DEL}}^{R^-}(u_0; u_1, x) &= \exists y (\phi_{\text{DEL}}^U(u_0; x) \vee (\phi_{\text{DEL}}^R(u_0; y) \wedge \phi_{\text{DEL}}^S(u_0; y, u_1))) \\ \phi_{\text{INS}}^{R_{\text{DEL}}^-}(u_0; u_1, x) &= \exists y \exists z (\phi_{\text{INS}}^S(u_0; x, z) \wedge \phi_{\text{INS}}^S(u_0; y, u_1)) \\ \phi_{\text{DEL}}^{R_{\text{DEL}}^-}(u_0; u_1, x) &= \exists y \exists z (\phi_{\text{DEL}}^S(u_0; x, z) \wedge \phi_{\text{DEL}}^S(u_0; y, u_1))\end{aligned}$$

Lemmas 4.6 and 4.7 together yield Theorem 4.2. We now finally prove Theorem 4.1. For this we need the following adaption of Lemma 3.8 to  $\Delta$ -semantics.

**Lemma 4.8.** (a) *For every  $\Delta$ -DYNUCQ $^-$ -program there is an equivalent  $\Delta$ -DYNCQ $^-$ -program.*

(b) *For every  $\Delta$ -DYNFO $^\wedge$ -program there is an equivalent  $\Delta$ -DYNFO $^\wedge$ -program.*

*Proof.* The proof uses the idea from the corresponding Lemma 3.8 for absolute semantics. We prove (a) only. The construction for (b) is exactly the same.

Let  $\mathcal{P}$  be a  $\Delta$ -DYNUCQ $^-$ -program. As in Lemma 3.8 we construct two programs  $\mathcal{P}'$  and  $\mathcal{P}''$  equivalent to  $\mathcal{P}$  for domains of size at least two and domains of size one, respectively. The construction of  $\mathcal{P}'$  is exactly the same as the construction for absolute semantics. For the construction of  $\mathcal{P}''$ , Lemma 4.9 (see below) is used.

A  $\Delta$ -DYNCQ $^-$ -program  $\mathcal{P}'''$  is obtained from  $\mathcal{P}'$  and  $\mathcal{P}''$  by using a modification of the construction used for the cases (b) and (c) in Lemma 3.8.

In order to delegate the case distinction to the initialization mapping, we use an additional 0-ary relation symbol  $U$  to ensure that interpretations of relations  $R'' \in \tau''$  never change for domains of size at least two and, analogously, interpretations of relations  $R' \in \tau'$  never change for domains of size one.

To achieve this,  $U$  is interpreted by true if and only if the domain is of size at least two and the update formulas of  $\mathcal{P}'$  and  $\mathcal{P}''$  are slightly modified as follows.



Update formulas  $\phi_\delta^{R'^+}$  and  $\phi_\delta^{R'^-}$  of a relation symbol  $R' \in \tau'$  in program  $\mathcal{P}'$  are replaced in  $\mathcal{P}'''$  by  $\phi_\delta^{R'^+} \wedge U$  and  $\phi_\delta^{R'^-} \wedge U$ . Hence the interpretation of  $R'$  changes only for domains of size at least two.

Similarly, update formulas  $\phi_\delta^{R''^+}$  and  $\phi_\delta^{R''^-}$  of a relation symbol  $R'' \in \tau''$  in program  $\mathcal{P}''$  are replaced in  $\mathcal{P}'''$  by  $\phi_\delta^{R''^+} \wedge \neg U$  and  $\phi_\delta^{R''^-} \wedge \neg U$ . Hence the interpretation of  $R''$  changes only for domains of size one.

The initialization of relation symbols from  $\tau' \cup \tau'' \cup \{Q'''\}$  is as in Lemma 3.8, and  $U$  is initialized as true if and only if  $|D| = 1$ .  $\square$

**Lemma 4.9.** *Every query on a 0-ary database can be maintained by a  $\Delta$ -DYNPROPCQ-program.*

*Proof.* Let  $\tau_{\text{in}}$  be an input schema with 0-ary relation symbols  $A_1, \dots, A_k$ . Further let  $\mathcal{Q}_1, \dots, \mathcal{Q}_m$  be an enumeration of all  $m = 2^{2^k}$  many queries on  $\tau_{\text{in}}$ . As in Lemma 3.9 we show that all of them can be maintained by one  $\Delta$ -DYNPROPCQ-program  $\mathcal{P}$  with auxiliary schema  $\tau_{\text{aux}} = \{R_1, \dots, R_m\}$  maintaining  $\mathcal{Q}_i$  in  $R_i$ , for every  $i \in \{1, \dots, m\}$ .

Our goal is to re-use the program constructed in Lemma 3.9 and the translation

$$\begin{aligned}\phi_\delta^{R_i^+}(\vec{u}; \vec{x}) &= \phi_\delta^{R_i}(\vec{u}; \vec{x}) \wedge \neg R_i(\vec{x}) \\ \phi_\delta^{R_i^-}(\vec{u}; \vec{x}) &= \neg \phi_\delta^{R_i}(\vec{u}; \vec{x}) \wedge R_i(\vec{x})\end{aligned}$$

Yet  $\neg \phi_\delta^{R_i}$  does not yield a DYNPROPCQ-formulas immediately.

The idea to solve this issue is to use two dynamic programs  $\mathcal{P}^\wedge$  and  $\mathcal{P}^\vee$  that both maintain all queries  $\mathcal{Q}_i$  in their auxiliary relations. The program  $\mathcal{P}^\wedge$  will be the program from Lemma 3.9 whereas  $\mathcal{P}^\vee$  will be a DYNPROP-program whose update formulas are disjunctions of atoms. Then the update formulas of  $R_i$  in  $\mathcal{P}^\vee$  will be used for defining  $\phi_\delta^{R_i^-}$ .

We make this more precise. By Lemma 3.9 there is a DYNPROPCQ-program  $\mathcal{P}^\wedge$  over schema  $\tau^\wedge = \{R_1^\wedge, \dots, R_m^\wedge\}$  that maintains  $\mathcal{Q}_i$  in  $R_i^\wedge$  with conjunctive quantifier-free update formulas. Analogously a dynamic program  $\mathcal{P}^\vee$  over schema  $\tau^\vee = \{R_1^\vee, \dots, R_m^\vee\}$  can be constructed that maintains  $\mathcal{Q}_i$  in  $R_i^\vee$  with disjunctive quantifier-free update formulas.

Then the update formulas for  $R_i$  in  $\mathcal{P}$  are constructed as

$$\begin{aligned}\phi_\delta^{R_i^+} &= \phi_\delta^{R_i^\wedge}[\tau^\wedge \rightarrow \tau] \wedge \neg R(\vec{x}) \\ \phi_\delta^{R_i^-} &= \neg \phi_\delta^{R_i^\vee}[\tau^\vee \rightarrow \tau] \wedge R(\vec{x})\end{aligned}$$

where  $\phi_\delta^{R_i^\wedge}[\tau^\wedge \rightarrow \tau]$  is obtained from  $\phi_\delta^{R_i^\wedge}$  by replacing symbols  $S^\wedge \in \tau^\wedge$  by  $S \in \tau$ , and  $\phi_\delta^{R_i^\vee}[\tau^\vee \rightarrow \tau]$  is obtained from  $\phi_\delta^{R_i^\vee}$  by replacing symbols  $S^\vee \in \tau^\vee$  by  $S \in \tau$ .

Those update formulas can be easily written as conjunctions. Negations can be removed by Lemma 4.5.  $\square$

*Proof (of Theorem 4.1).* The equivalence of (a) and (b) as well as of (c) and (d) follows from Lemma 4.5. Statements (a) and (c) are equivalent by Lemma 4.8. Further, (a) and (e) are equivalent by definition. The equivalence of (e) and (f) follows immediately by combining Lemmas 4.6 and 4.7 with Theorem 3.1.  $\square$

## 5 A dynamic characterization of first-order logic

In this section we characterize first-order queries as the class of queries maintainable by non-recursive UCQ<sup>-</sup>-programs and, equivalently, by non-recursive DYN $\exists^1$ FO-programs. Here  $\exists^1$ FO is the class of queries expressible by first-order formulas in prenex normal form with at most one existential quantifier and no universal quantifiers, and “non-recursive” is explained next.

A dynamic program is *non-recursive* if it has an acyclic dependency graph (as a directed graph). For every class  $\mathcal{C}$ , *non-recursive* DYN $\mathcal{C}$  refers to the set of queries that can be maintained by non-recursive DYN $\mathcal{C}$ -programs.

The objective of this section is to prove the following theorem.

**Theorem 5.1.** *For every query  $Q$  the following statements are equivalent*

- (a)  $Q$  can be expressed in FO.
- (b)  $Q$  can be maintained in non-recursive DYNFO.
- (c)  $Q$  can be maintained in non-recursive DYN $\exists^1$ FO.
- (d)  $Q$  can be maintained in non-recursive DYN $\forall^1$ FO.

With respect to the number of quantifiers in update formulas this result is optimal because the first-order definable alternating reachability query on graphs of bounded diameter cannot be maintained with quantifier-free update formulas [9]. Theorem 5.1 should be compared with the result of [9] that all  $\exists^*$ FO queries can be maintained in DYNQF.

Combining Theorem 5.1 with Theorem 3.1 immediately yields the following corollary.

**Corollary 5.2.** *Every first-order query can be maintained in DYN $\text{CQ}^-$ .*

The rest of this section is devoted to the proof of Theorem 5.1, more precisely to the equivalence of statements (a)-(c). The equivalence with (d) follows from Theorem 3.1 and the fact that its proof does not introduce recursion when applied to a non-recursive program.<sup>13</sup> It is obvious that (c) implies (b). For ease of presentation, we prove the remaining directions (a) $\Rightarrow$ (c) and (b) $\Rightarrow$ (a) for the input schema  $\tau_{\text{in}} = \{E\}$  where  $E$  is a binary relation symbol. The proofs can be easily adapted to general (relational) signatures.

The proof for (a) $\Rightarrow$ (c) makes use of the following normal form for FO. A formula  $\varphi$  is in *existential prefix form* if it has a prefix over  $((\neg\exists)|\exists)^*$  and no

<sup>13</sup>Alternatively, the proof of (a) $\Rightarrow$ (c) can be easily adapted to show (a) $\Rightarrow$ (d)

quantifier occurs after this prefix (e.g.  $\exists x \neg \exists y \neg (E(x, x) \rightarrow E(x, y))$ ) is in existential prefix form with prefix  $\exists \neg \exists$ . A formula in prefix normal form can be easily translated into existential prefix form by duality of universal and existential quantifiers. The *prefix length* of a formula in existential normal form is the number of existential and  $\neg$ -symbols in the maximal prefix ending with  $\exists$ .

The following example outlines the idea of the construction for the proof of (a) $\Rightarrow$ (c).

**Example 7.** Consider the query  $\mathcal{Q}$  defined by

$$\begin{aligned}\varphi &= \exists x \forall y (E(x, x) \rightarrow E(x, y)) \\ &\equiv \exists x \neg \exists y \neg (E(x, x) \rightarrow E(x, y))\end{aligned}$$

We construct a non-recursive dynamic DYN $\exists^1$ FO-program  $\mathcal{P}$  that maintains  $\mathcal{Q}$  under deletions only (for simplicity). The construction of  $\mathcal{P}$  applies the squirrel technique from Subsection 3.1. It uses a separate auxiliary relation  $R_\psi$  for each subformula  $\psi$  obtained from  $\varphi$  by stripping off a “quantifier prefix” from the existential prefix form of  $\varphi$ . The relation  $R_\psi$  reflects the possible states after a sequence of changes whose length equals the number of stripped off  $\neg$ - and  $\exists$ -symbols.

In order to update the query relation after the deletion of an edge, we maintain an auxiliary ternary relation<sup>14</sup>  $R_1$  that contains the result of the query  $\psi_1 \stackrel{\text{def}}{=} \neg \exists y \neg (E(x, x) \rightarrow E(x, y))$  for every choice  $a_1$  for  $x$  and every (possibly deleted) edge  $\vec{e}_1$ , that is  $(a_1, \vec{e}_1) \in R_1$  if and only if

$$(V, E \setminus \{\vec{e}_1\}, \{x \mapsto a_1\}) \models \neg \exists y \neg (E(x, x) \rightarrow E(x, y)).$$

Then we can define  $\phi_{\text{DEL}}^Q(\vec{v}_1) \stackrel{\text{def}}{=} \exists x R_1(x, \vec{v}_1)$  and it only remains to find a way to update the relation  $R_1$ . To this end, we maintain a further relation  $R_2$  that contains the result of  $\psi_2 \stackrel{\text{def}}{=} \exists y \neg (E(x, x) \rightarrow E(x, y))$  for every choice  $a_1$  for  $x$  and all (possibly deleted) edges  $\vec{e}_1, \vec{e}_2$ , that is  $(a_1, \vec{e}_1, \vec{e}_2) \in R_2$  if and only if

$$(V, E \setminus \{\vec{e}_1, \vec{e}_2\}, \{x \mapsto a_1\}) \models \exists y \neg (E(x, x) \rightarrow E(x, y)).$$

Then  $\phi_{\text{DEL}}^{R_1}(\vec{v}_1; x, \vec{v}_2) \stackrel{\text{def}}{=} \neg R_2(x, \vec{v}_1, \vec{v}_2)$  and it remains to update the relation  $R_2$ . Therefore we maintain a relation  $R_3$  that contains the result of  $\psi_3 = \neg(E(x, x) \rightarrow E(x, y))$  for every choice  $a_1, a_2$  for  $x, y$  and all (possibly deleted) edges  $\vec{e}_1, \vec{e}_2, \vec{e}_3$ . Then

$$\phi_{\text{DEL}}^{R_2}(\vec{v}_1; x, \vec{v}_2, \vec{v}_3) \stackrel{\text{def}}{=} \exists y R_3(x, y, \vec{v}_1, \vec{v}_2, \vec{v}_3)$$

and it remains to update relation  $R_3$  via

$$\phi_{\text{DEL}}^{R_3}(\vec{v}_1; x, y, \vec{v}_2, \vec{v}_3, \vec{v}_4) \stackrel{\text{def}}{=} \neg(E'(x, x, \vec{v}_1, \dots, \vec{v}_4) \rightarrow E'(x, y, \vec{v}_1, \dots, \vec{v}_4))$$

where  $E'$  is the edge relation obtained from  $E$  by deleting  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  and  $\vec{v}_4$ , that is  $E'(x, y, \vec{v}_1, \dots, \vec{v}_4)$  can be replaced by

$$E(x, y) \wedge (x, y) \neq \vec{v}_1 \wedge \dots \wedge (x, y) \neq \vec{v}_4.$$

---

<sup>14</sup>For simplicity we write  $R_1$  instead of  $R_{\psi_1}$ .

This completes the description of the program  $\mathcal{P}$  for  $\varphi$  which is easily seen to be non-recursive.

The following definition will be useful in both remaining proofs. For every first-order formula  $\varphi$  with  $k$  free variables and every sequence  $\delta = \delta_1 \dots \delta_j$  over  $\{\text{INS}, \text{DEL}\}$  let  $\varphi_{\delta_1 \dots \delta_j}^E$  be a  $(k + 2j)$ -ary formula such that for every graph  $G = (V, E)$ , every  $\vec{a} \in V^k$  and every instantiation  $\alpha = \delta_1(\vec{e}_1) \dots \delta_j(\vec{e}_j)$  of  $\delta$  with tuples  $\vec{e}_1, \dots, \vec{e}_j \in V^2$ :

$$\alpha(G) \models \varphi \text{ if and only if } G \models \varphi_{\delta_1 \dots \delta_j}^E(\vec{a}, \vec{e}_1, \dots, \vec{e}_j).$$

It is straightforward to construct  $\varphi_{\delta_1 \dots \delta_j}^E$ . It should be noted that  $\varphi_{\delta_1 \dots \delta_j}^E$  can be constructed such that its quantifier-prefix is the same as for  $\varphi$ . In particular, if  $\varphi$  is quantifier-free then  $\varphi_{\delta_1 \dots \delta_j}^E$  can be constructed quantifier-free as well. For example, if  $\delta = \text{INS DEL}$  and  $\varphi(\vec{x}) = \neg E(\vec{x})$  then

$$\varphi_{\text{INS DEL}}^E(\vec{x}, \vec{u}_1, \vec{u}_2) = \neg \left( (E(\vec{x}) \vee \vec{x} = \vec{u}_1) \wedge \neg(\vec{x} = \vec{u}_2) \right).$$

**Lemma 5.3.** *If a query is definable in FO, then it can be maintained in non-recursive  $\text{DYN}\exists^1\text{FO}$ .*

*Proof.* Inductively over the length of the prefix of a formula  $\varphi$  in existential prefix form, we prove that, for every finite sequence  $\delta_1 \dots \delta_j$ , the query defined by  $\varphi_{\delta_1 \dots \delta_j}$  is maintainable in non-recursive  $\text{DYN}\exists^1\text{FO}$ . The claim follows by setting  $j = 0$ . We construct dynamic programs where the result of the query defined by  $\varphi_{\delta_1 \dots \delta_j}$  is stored in the relation  $R_{\delta_1 \dots \delta_j}^\varphi$ .

For a formula  $\varphi$  with a prefix of length 0 (i.e. a quantifier-free formula), we define

$$\phi_{\delta_0}^{R_{\delta_1 \dots \delta_j}^\varphi}(\vec{v}_0; \vec{y}, \vec{v}_1, \dots, \vec{v}_j) \stackrel{\text{def}}{=} \varphi_{\delta_0 \dots \delta_j}^E(\vec{y}, \vec{v}_0, \dots, \vec{v}_j)$$

where  $\varphi_{\delta_0 \dots \delta_j}^E$  is as defined above (in the quantifier-free case).

For the induction step, let  $\varphi$  be a formula of prefix length  $i$ . By induction hypothesis, every query defined by  $\psi_{\delta_1 \dots \delta_j}$  where  $\psi$  has prefix length  $i - 1$ , can be maintained in non-recursive  $\text{DYN}\exists^1\text{FO}$ , for every sequence  $\delta_1 \dots \delta_j$  of modifications.

We distinguish the two cases  $\varphi(\vec{y}) = \exists x \psi(x, \vec{y})$  and  $\varphi(\vec{y}) = \neg \gamma(\vec{y})$ . If  $\varphi(\vec{y}) = \exists x \psi(x, \vec{y})$  then the dynamic program for  $\varphi$  and  $\delta_1 \dots \delta_j$  has auxiliary relations  $R_{\delta_0 \dots \delta_j}^\psi$  for  $\delta_0 \in \{\text{INS}, \text{DEL}\}$  containing the result of the query  $\psi_{\delta_0 \dots \delta_j}$ . Further,

$$\phi_{\delta_0}^{R_{\delta_1 \dots \delta_j}^\psi}(\vec{v}_0; \vec{y}, \vec{v}_1, \dots, \vec{v}_j) \stackrel{\text{def}}{=} \exists x R_{\delta_0 \dots \delta_j}^\psi(x, \vec{y}, \vec{v}_0, \dots, \vec{v}_j).$$

If  $\varphi(\vec{y}) = \neg \gamma(\vec{y})$  then the dynamic program for  $\varphi$  and  $\delta_1 \dots \delta_j$  has auxiliary relations  $R_{\delta_0 \dots \delta_j}^\gamma$  for  $\delta_0 \in \{\text{INS}, \text{DEL}\}$  containing the result of the query  $\gamma_{\delta_0 \dots \delta_j}$ . Further,

$$\phi_{\delta_0}^{R_{\delta_1 \dots \delta_j}^\gamma}(\vec{v}_0; \vec{y}, \vec{v}_1, \dots, \vec{v}_j) \stackrel{\text{def}}{=} \neg R_{\delta_0 \dots \delta_j}^\gamma(\vec{y}, \vec{v}_0, \dots, \vec{v}_j).$$

This yields a non-recursive  $\exists^1\text{FO}$ -program, for every  $\varphi_{\delta_1 \dots \delta_j}$ .  $\square$

We now turn towards proving the implication (b) $\Rightarrow$ (a) in Theorem 5.1. The following notion will be useful. A *topological sorting* of a graph  $(V, E)$  is a sequence  $v_1, \dots, v_n$  such that every vertex from  $V$  occurs exactly once and  $i > j$  for all edges  $(v_i, v_j) \in E$ . Every acyclic graph has a topological sorting. In particular, if  $R_1, \dots, R_m$  is a topological sorting of the dependency graph of a non-recursive dynamic program  $\mathcal{P} = (P, \text{INIT}, Q)$  then update formulas for  $R_1$  do only contain relation symbols from  $\tau_{\text{in}}$ . Further we can assume, without loss of generality, that  $R_m = Q$ . We say that  $R_i$  is on the  $i$ th level of the dependency graph.

**Lemma 5.4.** *If a query can be maintained in non-recursive DYNFO, then it can be expressed in FO.*

*Proof.* Consider a non-recursive dynamic DYNFO-program  $\mathcal{P} = (P, \text{INIT}, Q)$  over input schema  $\{E\}$ .

We start with some intuition. Let  $R$  be an auxiliary relation of  $\mathcal{P}$  which is (for simplicity) on the first layer of the topological sorting of the dependency graph of  $\mathcal{P}$ . That is, the update formulas  $\phi_{\text{INSE}}^R$  and  $\phi_{\text{DELE}}^R$  of  $R$  depend on the input relations only. There is no a priori upper bound on the complexity of the initialization process for  $R$ . However, after one modification step the relation is redefined via one of the first-order update formulas  $\phi_{\text{INSE}}^R$  or  $\phi_{\text{DELE}}^R$  which only use atoms over the input relations. Similarly, the auxiliary relations on higher levels of the dependency graph depend in a first-order fashion from the input structure after a constant number of modification steps. This is exploited in the proof.

More technically, the proof idea is as follows. For every modification pattern  $\delta = \delta_1 \dots \delta_j$  and every auxiliary relation  $R$ , a first-order formula  $\varphi_\delta^R$  is constructed that “precomputes” the state of  $R$  for every possible modification sequence with the pattern  $\delta$ . Thanks to non-recursiveness, once  $\delta$  is longer than the number of auxiliary relations, the formula  $\varphi_\delta^R$  can only use relations from the input schema. That is, it is just a first-order formula over  $\tau_{\text{in}}$ . We get the desired first-order formula for  $Q$  by choosing in  $\varphi_\delta^R$  a sufficiently long modification sequence  $\delta$  (by repeatedly inserting and deleting the same tuple).

We make this more precise now. Let  $\mathcal{Q}$  be a query which can be maintained by a non-recursive DYNFO-program  $\mathcal{P} = (P, \text{INIT}, Q)$  over schema  $\tau = \tau_{\text{in}} \cup \tau_{\text{aux}}$ . We assume for simplicity that  $\tau_{\text{in}} = \{E\}$ , for a binary symbol  $E$ . We let  $R_0 \stackrel{\text{def}}{=} E$  and assume that the auxiliary relations  $R_1, \dots, R_m$  are enumerated with respect to a topological sorting of the dependency graph of  $\mathcal{P}$  with  $R_m = Q$ .

We define inductively, by  $i$ , for every sequence  $\delta_1 \dots \delta_j$  with  $j \geq i$ , first-order formulas  $\varphi_{\delta_1 \dots \delta_j}^{R_i}(\vec{y}, \vec{x}_1, \dots, \vec{x}_j)$  over schema  $\tau_{\text{in}} = \{E\}$  such that  $\varphi_{\delta_1 \dots \delta_j}^{R_i}$  defines  $R_i$  after modifications  $\delta_1(\vec{x}_1) \dots \delta_j(\vec{x}_j)$ . More precisely  $\varphi_{\delta_1 \dots \delta_j}^{R_i}$  will be defined such that for every state  $\mathcal{S} = (V, E^{\mathcal{S}}, \mathcal{A}^{\mathcal{S}})$  of  $\mathcal{P}$  and every sequence  $\delta = \delta_1(\vec{a}_1) \dots \delta_j(\vec{a}_j)$  of modifications the following holds:

$$\mathcal{P}_\delta(\mathcal{S}) \upharpoonright R_i = \{\vec{b} \mid (V, E) \models \varphi_{\delta_1 \dots \delta_j}^{R_i}(\vec{b}, \vec{a}_1, \dots, \vec{a}_j)\} \quad (5)$$

Here  $\mathcal{P}_\delta(\mathcal{S}) \upharpoonright R_i$  denotes the relation stored in  $R_i$  in state  $\mathcal{P}_\delta(\mathcal{S})$ . For  $R_0 = E$  the formula  $\varphi_{\delta_1 \dots \delta_j}^E$  is as defined before the previous lemma. For  $R_i$  with  $i \geq 1$  the formula  $\varphi_{\delta_1 \dots \delta_j}^{R_i}(\vec{y}, \vec{x}_1, \dots, \vec{x}_j; \vec{y})$  is obtained from the update formula  $\phi_{\delta_j}^{R_i}(\vec{x}_j; \vec{y})$  of  $R_i$  by substituting all occurrences of  $R_{i'}(\vec{z})$  by  $\varphi_{\delta_1 \dots \delta_{j-1}}^{R_{i'}}(\vec{x}_1, \dots, \vec{x}_{j-1}, \vec{z})$  for all  $i' < i$ . Using induction over  $i$ , one can prove that the formulas  $\varphi_{\delta_1 \dots \delta_j}^{R_i}$  satisfy Equation 5. As  $\mathcal{P}$  is non-recursive, each formula  $\varphi_{\delta_1 \dots \delta_j}^{R_i}$  with  $j \geq i$  is over schema  $\{E\}$ .

The first-order formula  $\varphi$  for  $\mathcal{Q}$  over schema  $\tau_{\text{in}} = \{E\}$  can be constructed as follows. The formula “guesses” a tuple  $\vec{a} \in E$ , deletes and inserts it  $m$  times and applies  $\varphi_{(\text{DEL INS})^m}^{R_m}$  to the result (which is identical to the current graph), or (for the case that  $E$  is empty) it guesses a tuple  $\vec{a} \notin E$ , inserts and deletes it  $m$  times and applies  $\varphi_{(\text{INS DEL})^m}^{R_m}$  to the result.

More precisely,  $\varphi$  for  $\mathcal{Q}$  is defined by

$$\begin{aligned} \varphi(\vec{y}) \stackrel{\text{def}}{=} & \exists \vec{x} \left( (E(\vec{x}) \wedge \varphi_{(\text{DEL INS})^m}^{R_m}(\vec{y}, \underbrace{\vec{x}, \vec{x}, \dots, \vec{x}}_{2m\text{-times}})) \right. \\ & \left. \vee (\neg E(\vec{x}) \wedge \varphi_{(\text{INS DEL})^m}^{R_m}(\vec{y}, \underbrace{\vec{x}, \vec{x}, \dots, \vec{x}}_{2m\text{-times}})) \right). \end{aligned}$$

□

## 6 Discussion and Future Work

In this work, we studied dynamic conjunctive queries. We have shown that, contrary to the static setting, many fragments collapse in the dynamic world. Further we proved that DYNCQ captures DYNQF which implies that DYNCQ is strictly larger than DYNPROP. Moreover a close connection between absolute semantics and  $\Delta$ -semantics for conjunctive queries has been established. These results were summarized in Figure 1. Finally, it has been shown that dynamic conjunctive queries with negations capture (static) first-order logic.

All results are for arbitrary initialization mappings. However, they also hold in the setting with first-order definable initialization mappings. They do not carry over when the initialization mapping and updates have to be definable in the same class.

Although we have a good picture of the relationship of the various fragments now, it remains open whether the remaining classes DYNCQ, DYNCQ<sup>+</sup> and DYNFO can be separated or collapsed.

In addition to untangling the remaining variations of conjunctive queries, the dynamic quantifier hierarchy and quantifier alternation hierarchy, respectively, deserve a closer look. Lemma 3.10 shows that in the dynamic setting the  $\Sigma_i$ - and  $\Pi_i$ -fragment of first-order logic coincide. Whether there is a strict  $\Sigma_i$ -hierarchy remains open. Furthermore, the equivalence of  $\exists^*\text{FO}$  with absolute and  $\Delta$ -semantics does not immediately translate to fragments of FO with alternating quantifiers (although one of the direction does, see Lemma 4.6).

Capturing first-order logic by dynamic conjunctive queries with negations does not immediately yield performance gains (since a first-order query with  $k$  quantifiers is translated to a dynamic DYNCCQ<sup>-</sup>-program of arity at least  $k$ ). In future work we plan to study whether the work that has been started here can be used to improve the performance of query maintenance.

## References

- [1] Ashok K. Chandra and Philip M. Merlin. Optimal implementation of conjunctive queries in relational data bases. In *Proceedings of the ninth annual ACM symposium on Theory of computing*, pages 77–90. ACM, 1977.
- [2] Guozhu Dong, Leonid Libkin, and Limsoon Wong. On impossibility of decremental recomputation of recursive queries in relational calculus and SQL. In Paolo Atzeni and Val Tannen, editors, *Database Programming Languages (DBPL-5), Proceedings of the Fifth International Workshop on Database Programming Languages, Gubbio, Umbria, Italy, 6-8 September 1995*, Electronic Workshops in Computing, page 7. Springer, 1995.
- [3] Guozhu Dong, Leonid Libkin, and Limsoon Wong. Incremental recomputation in local languages. *Inf. Comput.*, 181(2):88–98, 2003.
- [4] Guozhu Dong and Jianwen Su. First-order incremental evaluation of datalog queries. In Catriel Beeri, Atsushi Ohori, and Dennis Shasha, editors, *Database Programming Languages (DBPL-4), Proceedings of the Fourth International Workshop on Database Programming Languages - Object Models and Languages, Manhattan, New York City, USA, 30 August - 1 September 1993*, Workshops in Computing, pages 295–308. Springer, 1993.
- [5] Guozhu Dong and Jianwen Su. Deterministic FOIES are strictly weaker. *Ann. Math. Artif. Intell.*, 19(1-2):127–146, 1997.
- [6] Guozhu Dong and Jianwen Su. Arity bounds in first-order incremental evaluation and definition of polynomial time database queries. *J. Comput. Syst. Sci.*, 57(3):289–308, 1998.
- [7] Guozhu Dong and Rodney W. Topor. Incremental evaluation of datalog queries. In Joachim Biskup and Richard Hull, editors, *Database Theory - ICDT'92, 4th International Conference, Berlin, Germany, October 14-16, 1992, Proceedings*, volume 646 of *Lecture Notes in Computer Science*, pages 282–296. Springer, 1992.
- [8] Kousha Etessami. Dynamic tree isomorphism via first-order updates. In Alberto O. Mendelzon and Jan Paredaens, editors, *Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 1-3, 1998, Seattle, Washington, USA*, pages 235–243. ACM Press, 1998.

- [9] Wouter Gelade, Marcel Marquardt, and Thomas Schwentick. The dynamic complexity of formal languages. *ACM Trans. Comput. Log.*, 13(3):19, 2012.
- [10] Erich Grädel and Sebastian Siebertz. Dynamic definability. In Alin Deutsch, editor, *15th International Conference on Database Theory, ICDT '12, Berlin, Germany, March 26-29, 2012*, pages 236–248. ACM, 2012.
- [11] Ashish Gupta, Inderpal Singh Mumick, and Venkatramanan Siva Subrahmanian. Maintaining views incrementally. In *ACM SIGMOD Record*, volume 22, pages 157–166. ACM, 1993.
- [12] William Hesse. The dynamic complexity of transitive closure is in DynTC<sup>0</sup>. *Theor. Comput. Sci.*, 296(3):473–485, 2003.
- [13] William Hesse. *Dynamic Computational Complexity*. PhD thesis, University of Massachusetts Amherst, 2003.
- [14] Christoph Koch. Incremental query evaluation in a ring of databases. In Jan Paredaens and Dirk Van Gucht, editors, *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2010, June 6-11, 2010, Indianapolis, Indiana, USA*, pages 87–98. ACM, 2010.
- [15] Christoph Koch, Yanif Ahmad, Oliver Kennedy, Milos Nikolic, Andres Nötzli, Daniel Lupei, and Amir Shaikhha. DBToaster: higher-order delta processing for dynamic, frequently fresh views. *VLDB J.*, 23(2):253–278, 2014.
- [16] Sushant Patnaik and Neil Immerman. Dyn-FO: A parallel, dynamic complexity class. In Victor Vianu, editor, *Proceedings of the Thirteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, May 24-26, 1994, Minneapolis, Minnesota, USA*, pages 210–221. ACM Press, 1994.
- [17] Oded Shmueli and Alon Itai. Maintenance of views. In Beatrice Yor-mark, editor, *SIGMOD'84, Proceedings of Annual Meeting, Boston, Massachusetts, June 18-21, 1984*, pages 240–255. ACM Press, 1984.
- [18] Volker Weber and Thomas Schwentick. Dynamic complexity theory revisited. *Theory Comput. Syst.*, 40(4):355–377, 2007.
- [19] Thomas Zeume and Thomas Schwentick. On the quantifier-free dynamic complexity of reachability. In Krishnendu Chatterjee and Jiri Sgall, editors, *Mathematical Foundations of Computer Science 2013 - 38th International Symposium, MFCS 2013, Klosterneuburg, Austria, August 26-30, 2013. Proceedings*, volume 8087 of *Lecture Notes in Computer Science*, pages 837–848. Springer, 2013.



- [20] Thomas Zeume and Thomas Schwentick. Dynamic conjunctive queries. In Nicole Schweikardt, Vassilis Christophides, and Vincent Leroy, editors, *Proc. 17th International Conference on Database Theory (ICDT), Athens, Greece, March 24-28, 2014.*, pages 38–49. OpenProceedings.org, 2014.
- [21] Thomas Zeume and Thomas Schwentick. On the quantifier-free dynamic complexity of reachability. *Inf. Comput.*, 240:108–129, 2015.